A Characteristics Approach to Optimal Taxation and Tax-Driven Product Innovation*

Henrik Jacobsen Kleven, London School of Economics

Joel Slemrod, University of Michigan

New Version: September 2009

Abstract

Any tax system imposing selective commodity taxation must have procedures for assigning different goods to tax rate categories. Real-world tax legislation does this on the basis of observable characteristics, allowing the tax system to handle a constantly evolving set of available goods. We recast the theory of optimal taxation in the language of characteristics using the Gorman-Lancaster model of consumer behavior, and present a theory of tax-driven product innovation and optimal line drawing. The paper consists of two parts. The first part presents optimal tax rules showing that characteristics can be used to gauge optimal tax rates in an intuitive way: the closer two goods are in characteristics space, the greater their substitutability and the smaller the optimal tax rate differential. The second part starts from the observation that, whenever the number of tax instruments is finite, tax systems have to draw lines that define tax-rate regions in characteristics space. Such lines are associated with notches in tax liability as a function of characteristics, creating incentives to introduce new goods (i.e., new characteristics combinations) in order to reduce tax liability. New goods introduced this way are socially inferior to existing goods. Second-best optimal tax systems draw lines so as to avoid such tax-driven product innovations; only goods on the characteristics possibility frontier are allowed in the market. Hence, although the tax system is second-best, the set of goods produced is first-best given the demand for characteristics.

*We would like to thank Alan Auerbach, Tim Besley, Robin Boadway, Louis Kaplow, Wojciech Kopczuk, Dan Shaviro, David Weisbach, and Shlomo Yitzhaki for helpful discussions and comments. Contact information for authors: (1) Henrik J. Kleven, Department of Economics, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, United Kingdom. Email: h.j.kleven@lse.ac.uk. (2) Joel Slemrod, University of Michigan, Ross School of Business, 701 Tappan St., Ann Arbor, MI 48109-1234. Email address: jslemrod@umich.edu.

1 Introduction

According to the theory of second-best efficient commodity taxation, the optimal tax rate on any good depends on the Slutsky matrix of compensated demand derivatives with respect to the prices of all goods. More generally, the optimal tax pattern may also depend on distributional objectives and on the pattern of externality and internality generation across goods. In general, optimal tax theory prescribes a different tax on each good.

Whatever the reason for selective commodity taxation, a non-capricious tax system must have procedures for distinguishing among goods subject to different tax rates. Real-world tax systems do that by appealing to the *characteristics* of the commodities. For example, American states' retail sales taxes often exempt food but not restaurant meals, requiring the tax law to draw a line between the two categories. This is done by appealing to a set of characteristics of a restaurant meal, and the line can be fine such as when grocery stores sell pre-prepared meals that may or may not be eaten on the premises, or set up in-store salad bars. The retail sales tax in the Canadian province of Ontario exempts basic food items such as flour but applies to other processed foods such as chocolate bars, requiring lines to be drawn, including one that subjects to tax "biscuits or wafers specifically packaged and marketed to compete with chocolate bars." Several European countries provide a subsidy for certain kinds of consumer services (e.g., cleaning, gardening, and house repair) based on a Ramsey-type justification that such services compete with untaxed home production. This requires the classification of services eligible for the subsidy based on observable characteristics.

The prominent role of characteristics in commodity tax systems is due to several factors. First, using observable characteristics is a natural and intuitive way to distinguish among different goods, or different groups of goods, and assign them to tax-rate categories. The alternative that the theory implies—classifying goods according to compensated elasticities—is infeasible, both because these elasticities are notoriously difficult to estimate precisely and because they would not be intuitive to either policy makers, voters or consumers in the way that characteristics-based rules are. Second, a shared characteristic plausibly signals something about the relative substitutability of the goods, and so may serve as a more readily measurable indicator of the ideal, but not observable, determinants of the appropriate tax rate. Third, modern economies produce a vast amount of different goods, and the set of available goods is constantly evolving. If tax laws were specified literally in terms of goods and their associated elasticities, then there would be no natural way to assign a new good to a tax category and the law would have to be re-specified to explicitly deal with the new good. In contrast, a characteristics-based rule for assigning tax rates to goods naturally handles the creation of new goods by limiting the tax policy choice to which characteristic-based category the new good falls in.

The first objective of this paper is to reformulate optimal commodity tax theory in the language of characteristics so that it matches up more easily with real tax systems. To do so we make use of the idea developed by Gorman (1980) and Lancaster (1966, 1975) that there exists a mapping of each good into characteristics space, and that it is the characteristics of goods, not the goods themselves, that generate utility.[1] We formalize the relationship between characteristics, substitutability and optimal tax rates, which allows us to explore the notion that shared characteristics can be used to gauge substitutability and hence optimal tax rate differentials. We show that the closer two goods are in characteristics space, the smaller the optimal tax rate differential.

The second objective of the paper is to address an important aspect of reality that has been ignored by the literature on optimal taxation, namely *tax-driven product innovation*. By this term we refer to the creation of new products, i.e., new characteristics combinations, which are introduced in the market in response to the tax system. For example, the prevalence of salad bars and cafes inside supermarkets may be in part a response to the differential tax treatment of restaurant meals and food purchased in grocery stores. In developing countries that impose higher taxes on automobiles than on other types of vehicles, industries emerge that produce low-tax vehicles that share many characteristics with cars. For example, the preferential tax treatment of motorcycles in Indonesia led to the creation of a new type of motorcycle with three wheels and long benches at the back seating up to eight passengers—car-like but not so car-like as to be taxed as cars. When Chile imposed much higher taxes on cars than on panel trucks, the market soon offered a redesigned panel truck that featured glass windows instead of panels and upholstered seats in the back.[2]

In the standard optimal tax model, addressing the creation of new goods is not tractable, because a change in the set of available goods must be associated with a new utility function

[1] Although Gorman's paper did not appear in a journal until 1980, it was originally written in 1956 and therefore predates Lancaster's work.

[2] These examples are taken from Harberger (1995).

(with new arguments) and therefore a new optimal tax problem. In the Gorman-Lancaster approach, on the other hand, because the set of characteristics that consumers value is stable, the utility function is robust to the introduction of new goods and product innovation can then be incorporated into the optimal tax problem.

In general, product innovation can come in two forms. It can either come as the introduction of new characteristics *combinations* within an already feasible set of characteristics *possibilities*, or it can come as the introduction of new characteristics combinations facilitated by expansions of the characteristics possibility set. The second form reflects technological progress driven by research and development, and may be labeled technology-driven product innovation. The first form of product innovation, sometimes called product variation or product variety, does not require a technological advance per se, and appears to be a ubiquitous and ongoing phenomenon among profit-seeking producers in the modern-day marketplace.[3] We focus on this form of product innovation and study its relationship with the tax system. It is shown that non-uniform taxation may give rise to the creation of goods that are socially inferior in characteristics space, but which are privately optimal for tax avoidance purposes. This represents a distortion in the set of available goods, which is different from the demand and supply distortions typically considered by public finance economists. The paper investigates the implications of this type of distortion for the optimal design of a tax system.

Once we allow for the creation of new goods, it becomes clear that a tax system must include procedures for assigning potential (but currently non-existing) goods to tax categories. In principle, this calls for a separate tax rate associated with every possible point in characteristics space, which corresponds to assuming an infinite number of tax rate instruments. If, as is obviously reasonable, the number of instruments is restricted to be finite, a tax system has to define subsets in characteristics space that correspond to tax-rate regions. This is entirely consistent with much real-world tax legislation that defines tax categories by listing a number of observable characteristics, and places any given commodity into the category with which it shares a majority of its characteristics. This procedure is often labelled *line drawing*. Although line drawing is a ubiquitous issue in real-world tax systems and a controversial point of contention among tax lawyers, there is little economic analysis of the issue. Thus, a third objective of this

[3] Indeed, Chamberlin (1953, p.3) stresses that "products are not in fact 'given'; they are continuously changed—improved, deteriorated, or just made different—as an essential part of the market process." Our paper pursues the idea that one reason that products are "just made different" is taxation.

paper is to take a first step toward establishing a theory of optimal line drawing.

We emphasize in the paper that a "line" shares many attributes of a "notch" in tax schedules, which refers to a discontinuity in the function of how tax liability relates to the tax base. Indeed, a line creates a notch in characteristics space, because the tax liability changes discontinuously when the characteristics vector of a good crosses the statutory line. Note that, as long as a continuum of tax rates is administratively infeasible, notches in characteristic space are an unavoidable feature of tax systems, not an idiosyncrasy.

We show that line drawing may lead to the introduction of new goods that are *more* intensive in the *high-tax* characteristic than the original goods in the same tax region. By moving a good towards the high-tax characteristic, but not crossing the line to the higher-tax region, consumers are able to obtain more of the high-tax characteristic without incurring the additional tax liability associated with the high-tax good. This form of product innovation may occur in two different regions in characteristics space. One is marginal product shifting around the existing goods, whereby new goods that provide slightly more of the high-tax characteristic replace the original goods. These are the supermarkets that provide some restaurant-like characteristics by setting up in-store salad bars. New goods introduced in this way are only slightly socially inferior to existing goods. The other is the introduction of new products exactly on the line that defines the border to the higher-tax region. These are the car-like motorcycles in Indonesia and the car-like panel trucks in Chile. Depending on the location of the line, such products may be very socially inferior to existing goods, but privately optimal as they deliver a large tax reduction by being located at the notch created by the line.

If the government can impose tax systems that includes any (finite) number of tax regions, it is always possible to design a non-uniform tax system that completely avoids tax-driven product innovation of socially inferior goods. We demonstrate that the second-best optimal tax system completely avoids the introduction of socially inferior goods; only goods located on the frontier of the no-tax characteristics possibility set are produced. The result can be seen as a form of the *production efficiency theorem* (Diamond and Mirrlees, 1971), although the model and reasoning is different from the standard setting. Like the classic production efficiency theorem, our result relies on strong assumptions regarding the commodity tax instruments possessed by the government, and should therefore be seem as an idealized theoretical benchmark.

Two remarks are worth making about this result. First, it does not rule out that the optimal tax system affects the set of goods in equilibrium, because the tax system will have substitution effects on the demand for characteristics, which may affect the derived demand for goods and lead to new goods being introduced or existing goods being eliminated. What the result implies is that any new good that arises due to such effects should be on a characteristics production frontier, such that the set of available goods under the second-best optimal policy is first-best efficient, conditional on demand. Second, the result also does not rule out that new goods are introduced as a result of technology-driven product innovations that allow previously infeasible characteristics combinations to be produced. Indeed, if new characteristics combinations are invented that expand the characteristics possibility set, our proposition implies that such characteristics combinations should be allowed by the second-best optimal tax system. Because such product innovations affect the underlying technology of the economy, it changes the parameters of the optimal tax problem and the tax system may have to undergo reform. But the new optimal tax system would satisfy the characterization that we provide in this paper.[4]

As far as we are aware, none of the earlier literature addresses the salient features of real-world tax systems that we explore: characteristics-based tax rules, tax-driven product innovation, and line drawing in characteristics space. Although we address these issues in the context of a Ramsey-style optimal consumption tax, we argue that they are a ubiquitous feature of all forms of taxation. This includes income taxation where different forms of income are treated differently, requiring lines to be drawn based on the characteristics of different income forms and where new types of compensation may be introduced in order to facilitate tax avoidance.

Our paper contributes to the large literature on optimal commodity taxation (e.g., Diamond and Mirrlees, 1971; Deaton, 1979, 1981; Christiansen, 1984; Saez, 2002, 2004), and proposes a framework that has implications for optimal income taxation and the theory of tax avoidance and evasion more generally.[5] Within the standard optimal commodity tax model, Gordon (1989), Weisbach (1999, 2000), Belan and Gauthier (2004, 2006), and Belan, Gauthier, and Laroque (2008) have studied a question related to line drawing: how to group goods into a limited set

[4] This discussion implicitly views technological progress as exogenous to tax policy. It is conceivable that true product innovations are endogenous to tax policy via effects on the amount and type of R&D. An analysis of optimal taxation under endogenous technical progress could in principle be incorporated into the framework we set out in this paper, and is an interesting topic for future research.

[5] For recent surveys of the literature on optimal commodity and income taxation, we refer to Auerbach and Hines (2002), Salanie (2003), Sørensen (2007), Kaplow (2008), Banks and Diamond (2008), and Crawford, Keen, and Smith (2008). The literature on tax avoidance and evasion has been surveyed by, e.g., Slemrod and Yitzhaki (2002) and Shaw, Slemrod, and Whiting (2008).

of tax categories. This set of papers offers rules for grouping goods based on compensated demand elasticities and possibly distributional weights. Related, Yitzhaki (1979) and Wilson (1989) analyze how to draw the line between a set of taxed and untaxed goods in a world where uniform taxation is optimal but where expanding the tax base is associated with administrative costs.

The rest of the paper is organized as follows. Section 2 sets out a characteristics approach to optimal taxation, and characterizes the optimal tax system assuming that the set of available goods is fixed. Section 3 allows for an endogenous set of available goods, and considers tax-driven product innovation and line drawing. Section 4 concludes.

2 A Characteristics Approach to Optimal Taxation

2.1 A Gorman-Lancaster Model

In this section we develop a characteristics approach to optimal taxation based on the theory of consumer behavior set out by Gorman (1980) and Lancaster (1966, 1975). The basic idea in the Gorman-Lancaster model is that goods are associated with characteristics, and that it is these characteristics that consumers value rather than the goods themselves. Any given good may be associated with many characteristics, and any given characteristic may be obtainable from several different goods. If we denote the quantity consumed of characteristics by $z_0, ..., z_M$, utility can be specified as

$$u = u\left(z_0, z_{1,} ..., z_M\right), \tag{1}$$

where characteristics are generated from goods $0, ..., N$ according to a consumption technology

$$z_k = c_{k0}x_0 + c_{k1}x_1 + ... + c_{kN}x_N, \quad k = 0, ..., M, \tag{2}$$

where c_{ki} is the amount of characteristic k contained in one unit of good i.[6]

This specification of the consumption technology makes two key assumptions. First, there is the assumption of *linearity* in characteristics generation. The basic idea is the following: if what we care about in a car is its fuel efficiency and its size, and if one car is characterized by a certain amount of fuel efficiency and a certain size, then a second identical car will have the same fuel efficiency and size. As the example suggests, the assumption of a linear mapping of goods

[6] We do not have to restrict the coefficients $c_{k0}, ..., c_{kN}$, nor the total amount of a characteristic z_k, to be positive. But we do assume that short sales of market goods are not possible, which implies $x_i \geq 0$, $\forall i$.

into characteristics space relies on the notion that characteristics are *intrinsic and objective*, and therefore reflect measurable features of a good that do not change with the amounts consumed of the different goods. Of course, there is still diminishing marginal returns to goods, but this effect operates solely through the utility function rather than through the mapping of goods in characteristics space. In section 2.5, we consider a generalized Gorman-Lancaster setup featuring nonlinear characteristics generation.

Second, there is *joint production* of characteristics because any given good x_i may produce several (and possibly all) characteristics. This jointness of characteristics is central to the Gorman-Lancaster approach and reflects the realistic idea that any single good will possess more than one valued characteristic.[7]

The budget constraint of the consumer is given by

$$\sum_{i=0}^{N} p_i x_i = 1, \tag{3}$$

where $p_0, ..., p_N$ denote the prices of goods to consumers and full income is normalized to be equal to 1. Producer prices are fixed and given by $q_0, ..., q_N$ such that the tax on good i equals $t_i = p_i - q_i$. We follow the convention in optimal tax theory by defining good 0 as leisure and goods $1, ..., N$ as market goods, assume that leisure cannot be taxed ($t_0 = 0$) and further normalize so that $q_0 = p_0 = 1$.

Notice that our specification assumes that tax liability is triggered by the purchase of goods, not the consumption of characteristics. The potential role of characteristics for taxation arises because the government may decide to set the tax rate on a given good to be a function of the characteristics of that good (and possibly all other goods). The optimal tax problem posed below explores the nature of the optimal relationship between characteristics and tax rates. The assumption that taxes apply to goods—but tax rates may depend on characteristics—is exactly consistent with actual tax legislation as discussed at the outset of the paper.

[7] The Gorman-Lancaster model is often put in the same category as the Becker (1965) model, which has been applied to optimal taxation by, e.g., Kleven (2000, 2004). However, besides the basic idea that market goods are not carriers of utility in themselves but enter into the production of utility-yielding commodities, the two models are fundamentally different. The Becker model deals with household production assuming that joint inputs (different market goods and time) are combined to produce a single output (a household activity). The Gorman-Lancaster model, on the other hand, considers the opposite situation where a single input (a market good) generates joint outputs (a bundle of characteristics). In other words, while the Gorman-Lancaster model considers a situation with fully joint production, the Becker model completely rules out joint production.

We may summarize the model in vector notation as follows[8]

$$u = u(\mathbf{z}), \quad \mathbf{z} = \mathbf{C}\mathbf{x}, \quad \mathbf{p}\mathbf{x} = 1, \tag{4}$$

where $\mathbf{C}$ is the $(M+1) \times (N+1)$ matrix of all characteristics coefficients. This matrix is assumed to have full rank, which amounts to an assumption that no two goods or characteristics (more precisely, linear combinations of goods or characteristics) are exactly identical.[9] Notice that, in the special case where $\mathbf{C}$ is diagonal, we may choose units so that $\mathbf{C} = \mathbf{I}$ and the model reduces to the standard model where $u = u(\mathbf{x})$.

An important feature of a Gorman-Lancaster model is the number of characteristics versus the number of goods i.e., the number of rows versus columns, in the consumption technology matrix. Three cases need to be distinguished:

1. The number of goods equals the number of characteristics, $N = M$. In this case, $\mathbf{C}$ can be inverted—there is a unique vector of goods associated with any given vector of characteristics. This implies that the consumer's problem can be formulated in two equivalent ways, either a goods formulation (maximizing $u(\mathbf{C}\mathbf{x})$ subject to $\mathbf{p}\mathbf{x} = 1$) or a characteristics formulation (maximizing $u(\mathbf{z})$ subject to $\mathbf{p}\mathbf{C}^{-1}\mathbf{z} = 1$) where $\mathbf{p}\mathbf{C}^{-1}$ is a vector of implicit prices on characteristics.

2. The number of goods is lower than the number of characteristics, $N < M$. In this case, $\mathbf{C}$ cannot be inverted, and it is no longer the case that any given characteristics vector can be obtained by appropriately selecting goods. This implies that the characteristics formulation of the consumer's problem is not feasible, and we therefore have to work with the goods formulation.

3. The number of goods is higher than the number of characteristics, $N > M$. Again, $\mathbf{C}$ cannot be inverted, but now any given characteristics combination can be obtained from more than one basket of goods. This implies that, at any given characteristics vector, the consumer chooses goods so as to minimize expenditures associated with obtaining those characteristics. With a linear consumption technology, expenditure minimization implies

[8] To simplify the notation, we do not specify if vectors/matrices are transposed or not.

[9] For example, if a column is linearly dependent on the other columns, this implies that one good is exactly identical in characteristics to another good (or a combination of other goods), and hence one of the two goods (whichever is more expensive) would never be purchased in equilibrium.

that each consumer will purchase at most as many goods as there are characteristics. Thus, with one representative consumer, the equilibrium cannot sustain more goods than characteristics, and this case then reduces to a case with $N^* \leq M$ where N^* is the number of cost-efficient goods associated with the M characteristics.

These remarks imply that, in a model with one representative consumer and a fixed set of goods, we may focus on situations where $N \leq M$ as in cases 1 and 2. When the set of goods is endogenized in Section 3, we allow for an unbounded number of potential goods (and hence $N > M$ as in case 3), and solve explicitly for the optimal set of goods in equilibrium.

In order to span all cases, we have to work with the goods formulation, i.e. $u(\mathbf{Cx})$ and prices $\mathbf{p}$. In this formulation, the consumer's first-order conditions can be written as

$$\boldsymbol{\nabla}\mathbf{u} \cdot \mathbf{C} = \lambda \mathbf{p} \qquad \text{or} \qquad \boldsymbol{\nabla}\mathbf{u} \cdot \mathbf{c}_i = \lambda p_i, \quad i = 0, ..., N, \tag{5}$$

where $\boldsymbol{\nabla}\mathbf{u} \equiv (u'_0, u'_1, ..., u'_M)$ is the gradient of the utility function with respect to characteristics, $\mathbf{c}_i = (c_{0i}, c_{1i}, ..., c_{Mi})$ is the ith column in $\mathbf{C}$ that reflects the characteristics provided by one unit of good i, and λ is the shadow price associated with the budget constraint. We may eliminate λ so as to emphasize the role of the marginal rates of substitution, which yields

$$MRS^x_{ij} \equiv \frac{\boldsymbol{\nabla}\mathbf{u} \cdot \mathbf{c}_i}{\boldsymbol{\nabla}\mathbf{u} \cdot \mathbf{c}_j} = \frac{p_i}{p_j}, \qquad i, j = 0, 1, ..., n, \tag{6}$$

where we write MRS^x_{ij} with a superscript x to emphasize that this is a MRS between market goods i and j, and not between characteristics. This MRS depends both on the properties of preferences as represented by $u(.)$ and on the mapping of $\mathbf{x}$ into $\mathbf{z}$ as captured by $\mathbf{C}$.

2.2 A Distance Function Approach

The two standard approaches to solving optimal commodity tax problems are the utility function (primal) approach and the expenditure function (dual) approach. These approaches yield optimal tax rules that, in general, depend on the entire Slutsky matrix of compensated demand derivatives of all goods with respect to all prices. Such rules do not lend themselves easily to simple and operational statements about tax policy without making strong simplifying assumptions about the structure of preferences. In order to understand the link between characteristics and optimal taxation using the standard approaches, one must first characterize the relationship between characteristics and compensated demand elasticities and then investigate the implications for the optimal tax rates as a function of elasticities. This is a very indirect and complicated way of analyzing the problem.

It turns out to be simpler to adopt a non-standard approach based on the *distance function* introduced into consumer theory by Gorman (1970, 1976) and first applied to optimal tax problems in two contributions by Deaton (1979, 1981).[10] The distance function approach leads to an optimal tax rule that depends, not on the entire substitution matrix, but only on the substitutability of different goods with leisure, where substitutability is defined in terms of Antonelli coefficients instead of Slutsky coefficients.[11] The Slutsky and Antonelli representations of the optimal tax system are equivalent, but the latter makes more straightforward the link between characteristics, substitutability, and optimal tax rates. In Section 2.4, we discuss the connection between our results and standard optimal tax results.

We define the distance function $a(\bar{u}, \mathbf{z}) = a(\bar{u}, \mathbf{Cx})$ as the scalar by which the characteristics vector $\mathbf{z}$ must be divided in order for the consumer to obtain an (arbitrary) utility level $\bar{u}$. That is, $a(\bar{u}, \mathbf{Cx})$ is implicitly defined by

$$u\left(\frac{\mathbf{z}}{a}\right) = u\left(\frac{\mathbf{Cx}}{a}\right) = \bar{u}, \tag{7}$$

where a "better" basket is associated with a higher value of a. Clearly, if $\bar{u}$ in (7) equals actual utility $u(\mathbf{Cx})$, then $a = 1$. We can then redefine the direct utility function $u = u(\mathbf{Cx})$ in terms of the distance function as

$$a(u, \mathbf{Cx}) = 1. \tag{8}$$

To find the derivative of the distance function with respect to x_i, we total differentiate eq. (7), which gives

$$\frac{\nabla \mathbf{u} \cdot \mathbf{c}_i}{a} - \frac{\nabla \mathbf{u} \cdot \mathbf{z}}{a^2} \cdot \frac{\partial a}{\partial x_i} = 0. \tag{9}$$

By evaluating (9) at the actual utility level ($a = 1$), and making use of the consumption technology ($\mathbf{z} = \mathbf{Cx}$), the first-order conditions (5), and the consumer budget ($\mathbf{px} = 1$), this can be rewritten as

$$a_i(u, \mathbf{Cx}) \equiv \frac{\partial a(u, \mathbf{Cx})}{\partial x_i} = \frac{\nabla \mathbf{u} \cdot \mathbf{c}_i}{\lambda} = p_i. \tag{10}$$

[10] The distance function is dual to the expenditure function, retaining its useful mathematical properties, but defined on primal variables (quantities consumed) instead of dual variables (prices).

[11] As shown by Deaton, it is possible to derive Antonelli representations of the optimal tax system based on the standard primal and dual approaches, but the distance function approach is much more direct.

Equation (10) states that, at an equilibrium, the first-order derivative of the distance function with respect to each good equals its price. Hence, the equation gives the price of each good as a function of demand $\mathbf{x}$ and utility u, and $(a_0, ..., a_N)$ are therefore the *inverse* compensated demand functions. From (10), we have

$$MRS_{ij}^{x}(u, \mathbf{Cx}) \equiv \frac{a_i(u, \mathbf{Cx})}{a_j(u, \mathbf{Cx})} = \frac{\boldsymbol{\nabla}\mathbf{u} \cdot \mathbf{c}_i}{\boldsymbol{\nabla}\mathbf{u} \cdot \mathbf{c}_j} = \frac{p_i}{p_j}, \tag{11}$$

so that a_i/a_j measures MRS_{ij}^{x} at a given utility level u, i.e. along an indifference surface. Moreover, we have

$$a_{ij}(u, \mathbf{Cx}) \equiv \frac{\partial a_i(u, \mathbf{Cx})}{\partial x_j} = \frac{\boldsymbol{\nabla}^2\mathbf{u} \cdot \mathbf{c}_i\mathbf{c}_j}{\lambda}, \tag{12}$$

where $\boldsymbol{\nabla}^2\mathbf{u}$ denotes the Hessian matrix of the utility function. The matrix of all the a_{ij}s is the Antonelli matrix, which is the generalized inverse of the Slutsky matrix.

Following Deaton (1979, 1981) and Deaton and Muellbauer (1980), we define goods i and j as complements if $a_{ij} > 0$, so that the marginal valuation of good i increases with the consumption of good j along an indifference surface. Conversely, if $a_{ij} < 0$, we say that goods i and j are substitutes. In the characterization of the optimal tax system, what will be particularly important is the relative complementarity of different goods with untaxed leisure (good 0). We say that, if $\frac{a_{i0}}{a_i} - \frac{a_{j0}}{a_j} = \partial \log(a_i/a_j)/\partial x_0$ is positive (negative), then good i is more complementary (substitutable) with leisure than is good j. Notice that these definitions of complementarity and substitutability based on Antonelli terms are not equivalent to those based on Slutsky terms.

2.3 Characteristics-Based Optimal Tax Rules

In this section, we express optimal tax rates directly as a function of Antonelli terms and demonstrate how they depend on characteristics. In solving the optimal tax problem, we will work with the distance function and use consumption levels (rather than tax rates) as control variables. Of course, by setting tax rates, the government is effectively controlling consumption levels.

In the optimal tax problem, the government faces a revenue constraint given by

$$\mathbf{tx} = R \quad \Leftrightarrow \quad (\mathbf{p} - \mathbf{q})\,\mathbf{x} = R. \tag{13}$$

The government chooses $\mathbf{x}$ in order to maximize utility $u(\mathbf{Cx})$ subject to the government budget constraint, the first-order conditions from the consumer's problem, and the zero tax on leisure.[12]

[12] Without this constraint, the optimal tax would be effectively lump-sum.

Using the distance function representation (in particular, eqs (8) and (10)), the optimal tax problem can then be stated as maximizing u with respect to $x_0, ..., x_N$, subject to

$$(i)\ a\left(u, \mathbf{Cx}\right) = 1, \quad (ii)\ \left(\mathbf{p} - \mathbf{q}\right)\mathbf{x} = R, \quad (iii)\ \mathbf{px} = 1, \quad (iv)\ \boldsymbol{\nabla}\mathbf{a} = \mathbf{p}, \tag{14}$$

where $\boldsymbol{\nabla}\mathbf{a} = (a_0, ..., a_N)$ is the gradient of the distance function with respect to $\mathbf{x}$, and where $p_0 = q_0 = 1$. By combining (ii)-(iii) and inserting $p_0 = q_0 = 1$, we may simplify the constraints in (14) as

$$(i)\ a\left(u, \mathbf{Cx}\right) = 1, \quad (ii')\ 1 - x_0 - \sum_{i=1}^{N} q_i x_i = R, \qquad (iv')\ a_0\left(u, \mathbf{Cx}\right) = 1. \tag{15}$$

Condition (ii') is a resource constraint for the economy. Condition (iv') includes only the first-order condition for good 0, because the conditions for goods $1, ..., N$ have become redundant as consumer prices $p_1, ..., p_N$ have been eliminated from the rest of the problem. The condition in (iv') implicitly defines x_0 as a function of utility u and the demand for all other goods, $x_1, ..., x_N$. We denote this function by $x_0\left(x_1, ..., x_N, u\right)$ and insert it into (i) and (ii'), so that the government is maximizing u with respect to $(x_1, ..., x_N)$ under (i)-(ii') and the relationship $x_0\left(.\right)$. The Lagrangian associated with this problem can be formulated as

$$\max_{x_1,...,x_N} \quad u - \rho\left[a\left(u, \mathbf{Cx}\right) - 1\right] + \mu\left[1 - x_0\left(.\right) - \sum_{i=1}^{N} q_i x_i - R\right], \tag{16}$$

where $\mathbf{x} = \left(x_0\left(.\right), x_1 ..., x_N\right)$ includes the function $x_0\left(.\right)$ as its first element.

The advantage of stating the optimal tax problem in this way is that it allows us to derive a simple and explicit solution for optimal tax rates as a function of characteristics. We can show the following:

Proposition 1 (Optimal Tax Rates) *The optimal tax rate differential on any pair of goods i and j is given by*

$$\frac{t_j}{p_j} - \frac{t_i}{p_i} = \frac{\rho + \mu}{\mu a_{00}} \cdot \frac{\partial \log\left(a_i/a_j\right)}{\partial x_0} = \frac{\rho + \mu}{\mu a_{00}} \cdot \frac{\partial \log\left[\sum_{k=0}^{M} \omega_{kj} \frac{c_{ki}}{c_{kj}}\right]}{\partial x_0}, \tag{17}$$

where $\omega_{kj} \equiv \frac{u'_k c_{kj}}{\boldsymbol{\nabla}\mathbf{u} \cdot \mathbf{c}_j}$ and $\sum_{k=0}^{M} \omega_{kj} = 1$.

Proof: In the appendix. $\square$

The first equality in eq. (17) does not exploit the structure imposed by the characteristics approach, and therefore corresponds qualitatively to a form one can obtain in a standard model without characteristics. It expresses the optimal tax rate differential between any pair of goods, i and j, in terms of the log-change in the marginal rate of substitution between i and j along an indifference surface as untaxed leisure varies. If good i is more complementary to leisure than good j so that $\partial \log(a_i/a_j)/\partial x_0 > 0$, then good i should be subject to a higher tax rate than good j.

The second equality in eq. (17) uses the characteristics structure to obtain a formula that shows how the optimal tax rate differential between two goods depends on their characteristics. To understand this expression, first note that the ω-parameters sum to 1 and therefore reflect a weighting of relative characteristics $\frac{c_{ki}}{c_{kj}}$ over $k = 0, ..., M$. Hence, the optimal tax rule expresses the tax rate differential between goods i and j in terms of the log-derivative (with respect to leisure) of a weighted average of relative characteristics $\frac{c_{ki}}{c_{kj}}$ over k. Because the characteristics coefficients themselves are fixed, this log-derivative reflects simply a re-weighting of relative characteristics. The impact of re-weighting relative characteristics is determined by the variation in relative characteristics, which in turn captures the distance between two goods in characteristics space. To see this, notice that being identical in characteristics does not require that the characteristics vectors be identical ($c_{ki} = c_{kj}\ \forall k$), but only that the two vectors are on the same ray in characteristics space ($c_{ki} = \gamma \cdot c_{kj}\ \forall k$), in which case there is no variation in $\frac{c_{ki}}{c_{kj}}$ over k.[13] Hence, the closer are two goods in characteristics space, the less variable is $\frac{c_{ki}}{c_{kj}}$ across characteristics k, in which case Proposition 1 tells us that the goods should be more equally taxed.

In order to illuminate what it means for goods to be close to one another in characteristics space, we next briefly consider a few specific cases. First of all, in the limit where goods i and j become identical in characteristics, i.e. where $\mathbf{c}_i = \gamma \cdot \mathbf{c}_j$, eq. (17) immediately implies that there be no differentiation in taxation, so that:

Corollary 1 (Uniform Tax Rates) *If a pair of goods, i and j, converge to one another in characteristics space, i.e. $\mathbf{c}_i \to \gamma \cdot \mathbf{c}_j$, then $\frac{t_j}{p_j} - \frac{t_i}{p_i} \to 0$.*

To see more clearly the relationship between characteristics and optimal tax differentiation

[13] In this case, the difference in characteristics is only a matter of different units (buying one unit of good i always gives the same characteristics as buying γ units of good j).

outside of the limiting special case of identical characteristics, consider an ordering of characteristics whereby $\frac{c_{ki}}{c_{kj}}$ is increasing in k. In this case, the more $\frac{c_{ki}}{c_{kj}}$ increases with k, the more different are the characteristics of two goods. The following proposition focuses on two special cases:

Proposition 2 (Optimal Tax Rates) *(i) Let $\frac{c_{ki}}{c_{kj}} = \gamma$ for $k = 0, ..., h$ and $\frac{c_{ki}}{c_{kj}} = \gamma + \Delta$ for $k = h+1, ..., M$, so that Δ (relative to γ) is a measure of the distance in characteristics between i and j. Then,*

$$\frac{t_j}{p_j} - \frac{t_i}{p_i} \propto \left\{ \frac{\sum_{k=h+1}^{M} \frac{\partial \omega_{kj}}{\partial x_0}}{\frac{\gamma}{\Delta} + \sum_{k=h+1}^{M} \omega_{kj}} \right\}. \tag{18}$$

(ii) Let $\frac{c_{ki}}{c_{kj}} = \gamma + \delta k$ for $k = 0, ..., M$, so that δ (relative to γ) is a measure of the distance in characteristics between i and j. Then,

$$\frac{t_j}{p_j} - \frac{t_i}{p_i} \propto \left\{ \frac{\sum_{k=0}^{M} \frac{\partial \omega_{kj}}{\partial x_0} \cdot k}{\frac{\gamma}{\delta} + \sum_{k=0}^{M} \omega_{kj} \cdot k} \right\}. \tag{19}$$

In both (18) and (19), the denominator is positive, and therefore $\frac{t_j}{p_j} - \frac{t_i}{p_i}$ is non-decreasing in absolute value in either measure of distance, Δ or δ. Except when the numerator is exactly zero (so that uniform taxation is optimal), $\frac{t_j}{p_j} - \frac{t_i}{p_i}$ is strictly increasing in absolute value in either Δ or δ.

Proof: Follows by inserting the assumptions into (17) and rearranging. □

These propositions demonstrate the intuitive notion that, as two goods diverge in characteristics space in an unambiguous way, the optimal tax rate differential increases in absolute value.

2.4 Connection to Standard Optimal Tax Rules

As an alternative to the Antonelli-based optimal tax rules presented above, it is possible to characterize the optimal tax system in terms of the more familiar Slutsky terms as in the standard Ramsey rule and its various specializations. Two points are worth noting about the connection of our characteristics-based optimal tax rules to standard optimal tax results.

First, although the Gorman-Lancaster model is based on the idea that preferences depend on characteristics rather than goods, it is possible to write utility as a function of the quantities consumed of goods by re-defining utility as $u(\mathbf{Cx}) \equiv \tilde{u}(\mathbf{x})$. This implies that all the standard rules that express the optimal tax system as a function of compensated demand elasticities can be established within a Gorman-Lancaster setting, but where the elasticities depend on the structure of the characteristics matrix $\mathbf{C}$. Intuitively, as two goods approach one another in characteristics space, they will become closer substitutes for one another and more equally substitutable for leisure, in which case standard optimal tax rules call for a smaller tax rate differential between the two goods.

Second, in the special case where the number of goods equals the number of characteristics, the optimal tax system has a particularly nice characterization.[14] In this case, the characteristics matrix $\mathbf{C}$ can be inverted, so that we can represent the model solely in terms of characteristics by writing the budget constraint as $\mathbf{p}^z\mathbf{z} = 1$, where $\mathbf{p}^z \equiv \mathbf{pC}^{-1}$ is a vector of implicit prices on characteristics. By setting tax rates on goods, the government is able to control implicit prices on characteristics, but the untaxability of leisure implies that one characteristics price cannot by controlled freely.[15] Under this formulation, it is possible to obtain all the familiar optimal tax rules, but where tax rates and elasticities pertain to characteristics rather than goods. This implies that the optimal tax rate on any given characteristic depends on its complementarity with untaxed leisure. Moreover, at any given optimal tax structure on characteristics, there is an implied optimal tax structure on goods. In particular, we want high tax rates on goods that are relatively intensive in the characteristics that are complementary to untaxed leisure. As two goods approach one another in characteristics space, they become more equally intensive in the characteristics complementary to leisure, and so the desired tax rate differential becomes smaller.

2.5 A Generalized Gorman-Lancaster Model

Like Gorman and Lancaster, we have focused on the case of a linear consumption technology linking goods and characteristics. Let us briefly consider a generalized specification allowing for a nonlinear generation of characteristics, i.e.

$$z_k = f_k(x_0, x_1, ..., x_N), \quad k = 0, ..., M. \tag{20}$$

[14] An earlier version of this paper worked through this example in detail.

[15] A particularly simple way of capturing this restriction is by adopting the convention that one characteristic—characteristic zero—*is* leisure, so that $z_0 \equiv x_0$ and $q_0^z = p_0^z = 1$.

The model is otherwise identical to the one set out earlier, and the optimal tax problem can be formulated in a way that is analogous to the problem in section 2.3. The optimal tax system can be represented as follows:

Proposition 3 (Optimal Tax Rates) *The optimal tax rate differential on any pair of goods i and j is given by*

$$\frac{t_j}{p_j} - \frac{t_i}{p_i} = \frac{\rho + \mu}{\mu a_{00}} \cdot \frac{\partial \log (a_i/a_j)}{\partial x_0} = \frac{\rho + \mu}{\mu a_{00}} \cdot \frac{\partial \log \left[\sum_{k=0}^{M} \omega_{kj} \cdot \frac{\partial f_k / \partial x_i}{\partial f_k / \partial x_j} \right]}{\partial x_0}, \tag{21}$$

where $\omega_{kj} \equiv \frac{u'_k \frac{\partial f_k}{\partial x_j}}{\sum_{k=0}^{M} u'_k \frac{\partial f_k}{\partial x_j}}$ *and* $\sum_{k=0}^{M} \omega_{kj} = 1$.

Proof: Analogous to the proof of Proposition 1. □

Proposition 3 is very similar to Proposition 1, except that the optimal tax rate differential now depends on $\frac{\partial f_k / \partial x_i}{\partial f_k / \partial x_j}$, the relative marginal generation of characteristics of goods i and j, where the marginal generation of characteristics is no longer constant but instead depends on what bundle of goods is chosen. Alternatively, we could label $\frac{\partial f_k / \partial x_i}{\partial f_k / \partial x_j}$ the "marginal rate of transformation" in the consumption technology. Because relative marginal characteristics are now endogenous, the log-derivative in the expression is not simply a re-weighting of fixed characteristics, but reflects also a change in the characteristics themselves. However, it is still the case that, *other things being equal*, the smaller the variation in marginal relative characteristics $\frac{\partial f_k / \partial x_i}{\partial f_k / \partial x_j}$ across k (the more identical goods i and j are on the margin), the smaller is the optimal tax rate differential, $\frac{t_j}{p_j} - \frac{t_i}{p_i}$, in absolute value.

3 Tax-Driven Product Innovation and Line Drawing

Recasting optimal taxation in characteristics space facilitates the modeling of an endogenous set of goods. Indeed, an important advantage of the fact that real-world tax legislation specifies commodity tax rates in terms of characteristics is its robustness to changes in the set of available goods. The creation of new goods is an important feature of modern economies that is ignored completely by optimal tax theory, in part because new goods are not easily tractable within the standard framework in which a new good implies a new utility function and hence a completely new optimal tax problem.

In addressing this issue, we return to the case of linear characteristics generation, because the creation of new goods poses the same conceptual problems in the nonlinear Gorman-Lancaster model as in the standard non-characteristics model, only shifted to a different level. In a nonlinear characteristics approach, a change in the set of available goods does not imply a new utility function, but it implies a new set of characteristics production functions and hence a new optimal tax problem. Indeed, the power of the Gorman-Lancaster approach as a tool of analysis relies crucially on the notion of linear characteristics generation.

In keeping with the tradition of optimal tax theory, the production side of the model we set out above was implicit and very simple: firms operate under perfect competition and convert labor into different goods using a linear constant-returns-to-scale production technology. This implies that producer prices $q_0, ..., q_N$ equal marginal costs, which are constant. A natural starting point for developing an optimal tax theory that incorporates the creation of new goods is to maintain this simplified view of production. Hence, we assume that firms can create new goods—i.e., new characteristics combinations—using a linear transformation of labor into goods and with no setup costs, implying that there are constant returns to scale in all goods. Moreover, free entry of firms ensures that firms have no market power in any good, new or old.[16]

New goods may be located at different points in characteristics space according to a technology that we specify below. The goods-generating technology is taken to be a primitive of the model, and therefore does not depend on the tax system. As discussed earlier, the analysis does not deal with the potential effect of taxation on technology-changing innovations that allow previously infeasible characteristics combinations to be produced ("technology-driven product innovation"). It deals instead with the effect of the tax system on product innovations that consist of a re-packaging of characteristics within an already feasible set in order to reduce tax liability ("tax-driven product innovation"). The evidence discussed in the introduction suggests that this is an empirically important phenomenon.

To clarify the central insights, we start by considering a special case where the consumer derives utility from only two characteristics (and leisure). The two-dimensional characteristics

[16] Product innovation raises interesting questions pertaining to increasing returns to scale (due to setup costs) and imperfect competition. Such issues are ignored in the standard optimal tax model (with a fixed set of available goods), and we also ignore them here. Our model should be interpreted as dealing with purely tax-driven product innovations rather than innovations motivated by gaining market power. Tax-driven innovations reflect a re-packaging of characteristics that require no technological innovation (no R&D), and are therefore not associated with significant setup costs.

model is helpful to establish intuition, because it allows a graphical exposition of the model and results. However, our main results do not rely on this simplification, and in section 3.4 we present an analysis of the general case.

3.1 A Two-Characteristics Model with Endogenous Goods

The setup described above implies that the full gain of introducing a new good accrues to the consumer, and we can then incorporate the choice of what goods are produced into the consumer's utility maximization problem. We pose the consumer's problem in two stages. First, we consider how the consumer optimizes the set of available goods and the demand for each good in the absence of a tax system. Then we introduce a tax system, and address how the consumer re-optimizes the set of goods produced and the demand for each good in the presence of taxes.

The consumer derives utility from leisure and two characteristics so that $u = u(x_0, z_1, z_2)$. Characteristics are generated from goods in a linear fashion, so that one unit of good i generates c_{1i} units of characteristic 1 and c_{2i} units of characteristic 2. As before, we denote by $\mathbf{c}_i \equiv (c_{1i}, c_{2i})$ the characteristics vector of good i. Unlike in the earlier model, consumers can choose how many goods will be produced and where in characteristics space they will be located within a set of feasible goods. We therefore have to allow for an arbitrary number of *possible* goods, and in fact we will allow for a continuum of possible goods. However, as mentioned earlier, a model with two characteristics can sustain at most two goods in equilibrium. If we start from a situation with two goods and a new good is introduced in the market, then, if the new good survives in equilibrium, it will replace one of the existing goods. Hence, although we allow for an unbounded number of potential goods, this is really a 2×2 model.

The linearity of the consumption technology implies that any producible good generates characteristics along a ray in characteristics space, as illustrated in Figure 1. The ray associated with good i, r_i, has a slope equal to the characteristics ratio of this good, $\frac{c_{2i}}{c_{1i}}$. Following Lancaster (1975), we will assume that goods can be put on the market on *any ray* in characteristics space and therefore with any ratio of characteristics. Certain characteristics ratios may be technologically difficult (costly) to produce, whereas other characteristics ratios may be easy (cheap) to produce. By choosing units of all goods so that producer prices equal one, if a given characteristics ratio is technologically difficult to achieve, this shows up in the feasible characteristics vector at the given producer price of one, not in the price itself. At any given ray r_i in characteristics space, there is a maximum obtainable level of characteristics (c_{1i}, c_{2i}) per unit of a good with the characteristics ratio along this ray. As shown in the figure, this implies a curve of producible characteristics combinations. This curve gives the maximum amount of characteristic 2 as a function of the amount of characteristic 1 per unit of a good. Notice that, in the absence of taxation and under the normalization of producer prices to be equal to one, this represents an *iso-cost* curve of producible goods. We label this the *Goods Possibility Frontier* (GPF).

The choice of goods in this model depends crucially on the shape of the GPF. Figure 1 depicts the GPF as a bumpy curve that contains both concave and convex portions. Consider first a concave segment such as the segment from point $\mathbf{c}_1$ (good 1) to point $\mathbf{c}_3$ (good 3). If goods 1 and 3 are put on the market, the consumer can obtain characteristics vectors $\mathbf{c}_1$ and $\mathbf{c}_3$ as well as any linear combination in between these two vectors.[17] However, no matter what characteristics bundle is consumed using goods 1 and 3 (say, point A), this bundle is strictly dominated by a bundle that can be achieved by introducing a single good at the appropriate ray in between goods 1 and 3 (say, good 2 on ray r_2). In general, under concavity of the GPF, it is better to introduce one good with the appropriate characteristics ratio than combining different goods to obtain the desired consumption of characteristics. Hence, if the GPF is *globally* concave, we observe only one good in equilibrium. The result that under global concavity only one good exists is not specific to the two-characteristic model, but extends to the case of M characteristics. The implication of this discussion is that global concavity cannot be an accurate depiction of the real world, which is one where consumers purchase a large number of different goods.

Consider instead a convex segment on the GPF such as the segment between points $\mathbf{c}_3$ (good 3) and $\mathbf{c}_5$ (good 5). Here we have the opposite situation of the one just described, because now a single good (such as good 4 at $\mathbf{c}_4$) is strictly dominated by a convex combination of the two goods on each side. Moreover, as we increase the distance between the two goods, we strictly expand the set of characteristics bundles that become possible through linear combinations of the two goods. In general, under *global* convexity, there is always an incentive to make goods more extreme by increasing how much they provide of one characteristic and reducing how much

[17] Only linear combinations *between* the two goods are obtainable given the assumption that "short sales" of goods are not feasible ($x_1, x_2 \geq 0$). This assumption is clearly reasonable in the context of consumption goods.

they provide of the other characteristic. There exists no optimum in this model, except if we impose bounds on the amount of characteristics that can be provided by goods, in which case the optimum consists of corner goods located at those bounds. For example, if it is not possible to produce goods that contain negative amounts of a given characteristic, each good will be located at a corner point where it contains the maximum amount of one characteristic and none of the other.[18] As before, this is a general point that extends to the M-characteristics case.[19] Hence, global convexity is also an unrealistic depiction of the real world, which features many goods that combine several characteristics at once.

Having ruled out both global concavity and global convexity as reasonable assumptions, it follows that the GPF is indeed a bumpy curve of the general form shown in the figure. The concave bumps on the curve are "natural" goods in the sense that they reflect characteristics combinations that are technologically easy to produce, i.e., they provide relatively large amounts of the two characteristics. The locally concave portions are therefore the natural candidates for the goods chosen in equilibrium.

To see what goods are chosen in equilibrium, Figure 2 uses the GPF to construct a *Characteristics Possibility Frontier* (CPF). The CPF shows the maximum amount of characteristic 2 as a function of characteristic 1 that can be consumed by combining different goods on the GPF.[20] For example, if the individual wants to consume z_1^* units of characteristic 1, the maximum amount of characteristic 2 is z_2^* and it is achieved by introducing goods 1 and 2 at $\mathbf{c}_1^*$ and $\mathbf{c}_2^*$. In general, the CPF consists of linear segments that are tangent to the GPF as well as curvy segments that follow the GPF around its concave peaks.[21] On each linear segment, characteristics bundles are generated by combining two goods, whereas on each curvy and strictly concave

[18] The assumption that negative characteristic generation is infeasible is just made for the purpose of this example. We do not make this assumption in the analysis (as shown in the figure), and it does not seem reasonable to do so. If negative characteristics are possible, and under global convexity, goods will be located at corner points with a positive amount of one characteristic and the largest possible negative amount of the other characteristic.

[19] Notice that, under global convexity and infeasibility of negative characteristic generation, goods are chosen in such a way that the characteristics matrix $\mathbf{C}$ becomes diagonal. As mentioned in Section 2, the model then reduces to a standard non-characteristics model. Hence, extending the Gorman-Lancaster model to allow for endogenous goods provides precise conditions under which the characteristics and standard approaches are equivalent.

[20] Because the GPF is an iso-cost curve of producible goods, the CPF is an iso-cost curve of consumption possibilities in characteristics space. In other words, the CPF is a budget line (accounting for endogenous goods) in characteristics space.

[21] Because the slopes of adjacent linear segments are different, they will not be tangent to the GPF at the same point. This is the reason that the CPF does not feature sharp kinks, but is instead characterized by "curvy kinks" around the local concave peaks.

segment bundles are generated from just one good.

As we have not yet introduced taxation into the model, and given that the model features no other market imperfections, the construction of the CPF in Figure 2 reflects both private and social efficiency. In particular, the set of goods that span the no-tax CPF corresponds to the set of both privately and socially efficient goods. There are many socially efficient goods (infinitely many, in fact), because we have continua of efficient goods around the concave peaks. However, at any given desired characteristics bundle, there will be at most two socially efficient goods associated with that particular bundle.

To make these remarks formal, we denote by $c_{2j} = g(c_{1j})$ the relationship between characteristics coefficients for an arbitrary good j on the GPF-curve. We can state the following:

Lemma 1 (Socially Efficient Goods) *With only two characteristics, the equilibrium consists of at most two goods. Assume that, in the absence of a tax system, goods* 1 *and* 2 *are introduced in the market and denote their characteristics vectors by* $\mathbf{c}_1^* \equiv (c_{11}^*, g(c_{11}^*))$ *and* $\mathbf{c}_2^* \equiv (c_{12}^*, g(c_{12}^*))$*. Goods* 1 *and* 2 *are the socially efficient goods and for any other producible good, j,with characteristics* $\mathbf{c}_j \equiv (c_{1j}, g(c_{1j}))$*, we have*

$$g(c_{1j}) \leq \nu^* \cdot g(c_{11}^*) + (1 - \nu^*) \cdot g(c_{12}^*), \tag{22}$$

where $\nu^* \equiv \frac{c_{12}^* - c_{1j}}{c_{12}^* - c_{11}^*}$.

Proof: In the appendix. $\square$

We order goods so that, in the no-tax equilibrium, goods 1 and 2 are put on the market.[22] The chosen varieties of these two goods, $\mathbf{c}_1^*$ and $\mathbf{c}_2^*$, depend on the properties of both $g(.)$ and $u(.)$.[23] We adopt the convention that varieties are chosen such that good 2 is the one relatively intensive in characteristic 2, i.e. $\frac{g(c_{12}^*)}{c_{12}^*} > \frac{g(c_{11}^*)}{c_{11}^*}$. All other potential goods are socially inferior to goods 1 and 2, meaning that the characteristics possibilities allowed by introducing any of these goods

[22] As mentioned, it is possible that only one good is introduced. Our analysis extends to this case, but the optimal commodity tax problem is obviously more interesting with (at least) two goods. In a later section where we generalize the analysis to the case of M characteristics and N goods, we will explicitly allow for $N < M$ in equilibrium.

[23] That is, the properties of $g(.)$ determine the set of goods (picked from the GPF) that maximize characteristics possibilities (as reflected by the CPF). In Figure 2, this set consists of point A, segment B to c_2^*, segment c_1^* to D, and point E (there are more goods in this set to the left of A and to the right of E). We refer to this as the set of goods that span the CPF. Then, at a given CPF, the properties of $u(.)$ determine which consumption bundle is chosen (e.g., z^* in Figure 2), and therefore determine the specific goods (e.g., c_1^* and c_2^*) that are chosen from among the entire set that spans the CPF.

are in the interior of the characteristics possibility set generated by goods 1 and 2. Eq. (22) in Lemma 1 gives the formal definition of a "socially inferior" good.

3.2 The Effect of a Tax System on the Set of Goods

We now consider a government that must raise revenue from commodity taxation. Because the no-tax equilibrium sustains only goods 1 and 2, it might seem that the tax system could simply specify a rate t_1 applying to goods on ray r_1, and a rate t_2 applying to goods on ray r_2. But such a policy would be incomplete because, if a new good were to be introduced along a new ray in characteristics space, the tax system would have no way of dealing with the new good. Hence, in order to be robust to the introduction of new goods, a tax system must specify tax rates associated with both existing *and potential* combinations in characteristics space. With an unconstrained (by administrative considerations) set of instruments, a tax system could specify a selective tax rate on each potential good in characteristics space. However, this implies an infinite number of distinct tax rates, which is obviously unrealistic. Instead, real-world tax systems define regions—draw lines—in characteristics space that are subject to different rates of tax. We start by considering a government that defines two tax regions, and therefore sets tax rates t_1 and t_2 along with a line separating the two regions.[24] The line is a ray in characteristics space with slope ℓ. If a good i is characterized by $\frac{c_{2i}}{c_{1i}} > \ell$, it is taxed at rate t_2; otherwise it is taxed at rate t_1. Unless uniform taxation is optimal, the optimal line has to be located between the characteristic rays of the two existing goods, i.e. $\frac{g\left(c_{12}^*\right)}{c_{12}^*} > \ell \geq \frac{g\left(c_{11}^*\right)}{c_{11}^*}$.

Figure 3 illustrates the implications of a non-uniform tax system (t_1, t_2, ℓ) that imposes a higher tax rate on goods that are relatively intensive in characteristic 2 ($t_2 > t_1$). In the absence of taxation, the Goods Possibility Frontier is represented by the dashed curve GPF*, and the Characteristics Possibility Frontier is given by the solid line CPF*. Note that CPF* is tangent to GPF* at points $\mathbf{c}_1^*$ and $\mathbf{c}_2^*$, which are the privately and socially optimal goods associated with characteristics bundles in the range between these goods. The introduction of the tax system shifts down the CPF associated with goods 1 and 2 to the solid line CPF_{12}^t. The tax system also shifts down the GPF, because it is an *iso-cost* curve of possible goods. The downward shift at any ray is proportional to the tax rate and is therefore different on each side of the line: the

[24] Later on, we consider the possibility that the government may want to define more than two tax regions and hence draw additional lines in characteristics space.

GPF shifts down according to t_2 on the left-side of the line, while it shifts down according to t_1 on the right-side. The after-tax GPF, labelled GPF^{t}, is associated with a discontinuity (notch) at the line ℓ, reflecting that the tax liability associated with a good changes discretely as its characteristics cross the statutory line.

As shown in the figure, some goods that were dominated by other goods in the absence of taxation are no longer so. New goods have become potentially optimal in three different regions. First, goods that are similar to good 1 but more intensive in characteristic 2 (the segment from E to F) are located above the after-tax CPF associated with goods 1 and 2, so that introducing such goods would expand consumption possibilities. These are low-tax goods that represent slightly modified versions of the original low-tax good with a little bit more of the high-tax characteristic.[25] It is a general result that, for a line strictly above ray r_1, marginal product shifting around the low-tax good is optimal. To see this, notice that a proportional shift of the GPF does not affect its slope along rays in characteristics space, and therefore the slopes of GPF^{t} and GPF^* are the same along ray r_1. Hence, because GPF^* is tangent to CPF^* at ray r_1 and CPF^* is steeper than CPF^t_{12}, we have that GPF^{t} must be steeper than CPF^t_{12} at ray r_1. This implies that there exists a range of producible goods close to ray r_1 that can expand the set of consumption possibilities. The intuition for this result has two components. By replacing the original low-tax good by a low-tax good that provides more of the high-tax characteristic per unit of the good, the consumer is able to obtain a given amount of the high-tax characteristic by buying more of the low-tax good and less of the high-tax good, and thereby reduce tax liability. Moreover, if the new good is close enough to the original good, it is only slightly inferior from a technological point of view. Indeed, by the envelope theorem, as we move the characteristics of a good marginally away from their optimal composition, there is no first-order loss to the consumer.

Second, goods that are almost the same as good 2 on its left side (on the segment between A to B) now dominate good 2 in characteristics space. These are high-tax goods very much like the original high-tax good, but that provide even more of the high-tax characteristic. As above, this is a general phenomenon that can be understood from the observations that the no-tax and after-tax GPFs have identical slopes along rays along with the fact that CPF^t_{12} is flatter than

[25] In the discussion, we refer to characteristic 2 as the "high-tax characteristic", although this is a somewhat imprecise terminology as taxes apply to goods (depending on their characteristics), not to characteristics per se.

CPF*, implying that GPF$^{\mathrm{t}}$ is steeper than CPF$^*_{12}$ around point B. The economic intuition for the optimality of shifting high-tax goods towards the high-tax characteristic is essentially the same as the intuition for shifting low-tax goods in the same direction: conditional on a good being on the high-tax side of the line, making it more extreme in the high-tax characteristic does not trigger additional tax liability per unit of the good, and allows the consumer to obtain a given total amount of the high-tax characteristic while buying fewer units of the high-tax good. Moreover, if the good is located close enough to the original high-tax good, it will be only slightly inferior from the technological viewpoint.

Third, depending on the location of the line, there may be new goods close to the line (at the segment from C to D) that are able to expand the set of characteristics possibilities. As Figure 3 is drawn, the good located exactly on the line dominates the other possible goods in this range, but this is not always true and depends on the shape of the GPF around the line. What makes the good exactly on the line particularly desirable in the figure is the fact that the line is located at a convex portion of the GPF. If the line was instead located at a locally concave portion of the GPF, we might see new products close to, but to the right of, the line. In general, if the line is sufficiently close to the original high-tax good, we always see tax-driven product innovation close to the line, on the low-tax side. In particular, as we move the line arbitrarily close to the original high-tax good, we allow the introduction of a good that is almost identical to the high-tax good, but on the low-tax side of the line, allowing the consumer to completely avoid the high tax rate and with no first-order loss from the new good being technologically inferior. As we will see later, this kind of tax system (rates plus line) is almost certainly not optimal.

It is theoretically possible that, in addition to product innovation around the original goods and around the line, a tax system may induce product innovation at interior local peaks in the low-tax region. This may be the case if the GPF features local peaks that, in the absence of a tax system, are dominated by surrounding higher peaks. If such peaks exist on the low-tax side of the line, there may be privately optimal goods in those regions once taxes are imposed.

To clarify the effect of a tax system on the choice of goods, Figure 4 illustrates the after-tax Characteristics Possibility Frontier, CPF^t, that accounts for the re-optimization of the set of goods in response to the tax system. In between points A and C, the after-tax CPF combines a new high-tax good (that is slightly more intensive in the high-tax characteristic) with a new low-tax good located on the line. In between points C and E, the after-tax CPF combines two new low-tax goods, the one right on the line and one close to the original low-tax good that is slightly more intensive in the high-tax characteristic. In between points E and F, characteristics are generated from just one good, a new low-tax good close to the original low-tax good. The goods chosen in equilibrium depend on the demand for characteristics and therefore on preferences.

These arguments imply that, as the government introduces non-uniform commodity taxation based on characteristics, privately optimal goods become *more* intensive in the *high-tax* characteristic within each tax region. At the same time, across-region substitution effects may lead goods in the high-tax region to disappear completely (e.g., if the consumer's optimum is at a point such as D in Figure 4). Whatever the set of goods in equilibrium, the total *consumption* of the high-tax characteristic will go down in response to the tax system (as the CPF^t is everywhere flatter than CPF^*), but the consumer is able to counteract this effect somewhat by re-packaging the characteristics of goods within each tax region to provide more of the high-tax characteristic.[26]

We can state the following:

Proposition 4 (Tax-Driven Product Innovation) *Consider the effects of a non-uniform tax system with two tax regions, i.e.,* (t_1, t_2, ℓ) *where* $t_2 \neq t_1$ *and* $\frac{g(c_{12}^*)}{c_{12}^*} > \ell \geq \frac{g(c_{11}^*)}{c_{11}^*}$. *We have:*

a. *The original high-tax good is always eliminated. Either high-tax goods disappear completely, or a new high-tax good appears that contains more of the high-tax characteristic than the original good.*

b. *The original low-tax good may survive or it may be eliminated. If it is eliminated, one or two new low-tax goods will appear that each contain more of the high-tax characteristic than the original good. New low-tax goods may be located either exactly on the line* ℓ *or at the interior of the low-tax region.*

Proof: In the appendix. □

Each possible type of tax-driven product innovation in the proposition is distortionary, because each represents a change from the set of socially efficient goods. The new goods are introduced in response to distorted price signals, and this behavioral response lowers tax revenue and

[26] The statement that the consumption of the high-tax characteristic goes down in response to taxation implicitly assumes that the uncompensated demand curve for that characteristis is negatively sloped.

creates deadweight loss. Note, however, that the different types of tax-driven product innovation may have very different welfare implications. In particular, new goods located close to the original goods are only marginally socially inefficient and therefore imply small welfare losses. By contrast, the introduction of new goods located close to the line cause larger welfare losses if the line is drawn at a place with very inefficient goods (as in Figure 4). Such goods may be optimal to the consumer despite being strongly socially inefficient because they reduce tax liability substantially, but they are associated with large welfare losses.

We note the following:

Remark 1 (Two Tax Regions) *A non-uniform tax system* (t_1, t_2, ℓ) *where* $t_2 > t_1$ *and* $\frac{g(c_{12}^*)}{c_{12}^*} > \ell \geq \frac{g(c_{11}^*)}{c_{11}^*}$ *inevitably causes tax-driven product innovation. Setting* $\ell = \frac{g(c_{11}^*)}{c_{11}^*}$ *avoids product innovation on the low-tax side of the line, but in this case there will still be product innovation on the high-tax side of the line.*

As background for our analysis of optimal tax systems in the next section, it is helpful to consider tax systems that define more than two tax regions. In particular, note that with three tax regions it is possible to eliminate tax-driven product innovation:

Remark 2 (Three Tax Regions) *Consider a non-uniform tax system* $(t_1, t_2, t_3, \ell_1, \ell_2)$ *where* $t_3 > t_2 > t_1$ *and* $\ell_2 \geq \frac{g(c_{12}^*)}{c_{12}^*} > \ell_1 \geq \frac{g(c_{11}^*)}{c_{11}^*}$. *At any given* t_1 *and* t_2, *this tax system can avoid tax-driven product innovation by setting* $\ell_1 = \frac{g(c_{11}^*)}{c_{11}^*}$, $\ell_2 = \frac{g(c_{12}^*)}{c_{12}^*}$, *and selecting* t_3 *sufficiently high. Under this tax system, the socially efficient goods survive in equilibrium and are both located at a border between tax regions. No good in equilibrium is subject to the high tax rate* t_3.

Hence, if the government has a sufficient number of commodity tax rates (i.e., tax regions), it is feasible to completely avoid the distortion associated with a change in the set of available goods by drawing lines appropriately. However, because of the inability to tax leisure, a commodity tax system that does not distort the set of available goods is still not a first-best tax system. In a second-best environment, it is not a priori obvious that the optimal tax system features exactly the same set of goods that is associated with the first-best allocation. After all, an important implication of the theory of second best is that, typically, it is not optimal to eliminate distortions completely: it is better to have several small distortions than to have large distortions somewhere and none elsewhere. The most famous and most consequential exception to this guideline is the *production efficiency theorem* by Diamond and Mirrlees (1971) showing that the optimal tax system maintains the economy on the production possibility frontier. The next section characterizes second-best optimal allocation in our setting, and demonstrates that it does in fact avoid the introduction of goods that are socially inefficient.

3.3 The Second-Best Optimal Allocation

We characterize the second-best optimal allocation, assuming that the government must raise a given revenue from commodity taxation and that it cannot tax leisure, as in the Ramsey optimal tax problem considered in Section 2. We will show that the solution to this problem does not feature the introduction of socially inferior goods, and that the set of goods in equilibrium is therefore first-best. In order to establish this, it is sufficient to allow for one new good, and show that no matter where this good is located within the feasible set, it is dominated by the socially efficient goods in the government's problem. This implies that we have to allow for three potential goods; the socially efficient goods 1 and 2 associated with characteristics $\mathbf{c}_1^*$ and $\mathbf{c}_2^*$ and a potential new good (say, good 3) associated with characteristics $(c_{13}, g(c_{13}))$ that is chosen by the consumer in light of the tax system.

In this section, we impose additional structure on the problem in order to simplify the analysis, while in the next section we demonstrate the result in a very general setting. In particular, it greatly simplifies the analysis to normalize the characteristics of the socially efficient goods, so that good 1 delivers only characteristic 1 and good 2 delivers only characteristic 2, i.e. $\mathbf{c}_1^* \equiv (1,0)$ and $\mathbf{c}_2^* \equiv (0,1)$. The Goods Possibility Frontier, $g(.)$, has to be such that these are indeed the socially efficient goods. From Lemma 1, this implies $g(c_{13}) \leq 1 - c_{13}$ for the socially inefficient good 3. This imposes a very simple structure on the problem, but as we show in Section 3.4, it involves no loss of generality.

The government chooses x_1, x_2, and x_3 so as to maximize the representative consumer's utility subject to the government revenue constraint, the consumer's first-order conditions, and the zero tax on leisure. Under the normalizations of good 1 and 2, the utility function is given by $u(x_0, x_1 + c_{13}x_3, x_2 + g(c_{13})x_3)$. As in Section 2, we solve the optimal tax problem by specifying it in terms of the distance function. The distance function can be written as

$a\left(\bar{u}, x_0, x_1 + c_{13}x_3, x_2 + g\left(c_{13}\right)x_3\right)$ and is implicitly defined by

$$u\left(\frac{x_0}{a}, \frac{x_1 + c_{13}x_3}{a}, \frac{x_2 + g\left(c_{13}\right)x_3}{a}\right) = \bar{u}, \tag{23}$$

so that the utility function can be redefined as $a\left(u, x_0, x_1 + c_{13}x_3, x_2 + g\left(c_{13}\right)x_3\right) = 1$. The government budget constraint can be combined with the consumer budget constraint into a resource constraint:

$$1 - x_0 - x_1 - x_2 - x_3 = R. \tag{24}$$

The government must also account for a condition reflecting that the consumer optimizes untaxed leisure at any given choice of x_1, x_2, and x_3. Thus,

$$\frac{\partial a\left(.\right)}{\partial x_0} = 1, \tag{25}$$

which implicitly defines a function $x_0 = x_0\left(u, x_1 + c_{13}x_3, x_2 + g\left(c_{13}\right)x_3\right)$. We may then formulate the government's problem as choosing x_1, x_2, and x_3 to maximize

$$u - \rho\left[a\left(u, x_0\left(.\right), x_1 + c_{13}x_3, x_2 + g\left(c_{13}\right)x_3\right) - 1\right] + \mu\left[1 - x_0\left(.\right) - x_1 - x_2 - x_3 - R\right]. \tag{26}$$

We can show the following:

Proposition 5 (Product Efficiency) *In the second-best optimal allocation, the socially inferior good* 3 *is not introduced. Only goods on the no-tax Characteristics Possibility Frontier may be produced and consumed. Thus, despite the allocation being second-best, the set of available goods is first-best* (*at any given demand for characteristics*).

Proof: In the appendix. □

The second-best optimal tax system completely avoids tax-driven product innovations of socially inferior goods, i.e. goods located in the interior of the no-tax Characteristics Possibility Frontier.

Four points are worth noting about this proposition. First, the result is related to the Diamond-Mirrlees production efficiency theorem, although the model and proof are different from their setting. While the classic result deals with the production of goods from inputs, Proposition 5 deals with the "production" of characteristics from goods. In other words, the proposition above is a statement about the choice of products, not the choice of inputs, and it might alternatively be labeled a *product* efficiency theorem.

Second, the result imposes no a priori restrictions on the set of tax instruments *besides the inability to tax leisure*. This is consistent with the Diamond-Mirrlees setting, which also assumes that there is no limit to the ability of the government to vary tax rates across goods. While Proposition 5 is a statement about allocation, not implementation, the implied tax policy can be inferred from Remarks 1 and 2. We see that, in order to implement the second-best optimal allocation with a tax policy, the government has to define at least three tax regions in two-dimensional characteristics space. If, for administrative reasons, the government is restricted to only two tax regions, i.e. (t_1, t_2, ℓ), the optimal instruments under such a system will necessarily involve some purely tax-driven product innovation. Moreover, for a three-region tax system to avoid tax-driven product innovation, lines must be located exactly at the socially optimal goods, requiring that the government can perfectly observe characteristics. While this is a strong assumption, it is not conceptually different from the canonical assumption in optimal tax theory that the government perfectly observes behavioral elasticities. In fact, as discussed earlier, because the characteristics coefficients in a Gorman-Lancaster model should be seen as objective and measurable entities that do not change with the quantities consumed of goods, they are likely to be more observable than price elasticities.

Third, while the proposition rules out socially inefficient goods in the second-best optimum, it does not rule out that the optimal tax system affects the set of goods in equilibrium. To see this, notice that any non-uniform tax system is associated with substitution effects on the amounts consumed of different characteristics, which may affect the derived demand for goods generating those characteristics. In particular, as demand shifts toward the low-tax characteristic, we may see the introduction of goods that are relatively intensive in the low-tax characteristic or the elimination of goods that are intensive in the high-tax characteristic. This is a form of tax-driven product innovation, but one that is driven by traditional substitution effects on consumer demand and, critically, any new goods introduced this way will be efficient, i.e. located on the no-tax Characteristics Possibility Frontier.

Fourth, the result also does not rule out that new goods are introduced as a result of technology-driven product innovations that allow previously infeasible characteristics combinations to be produced. Indeed, if new characteristics combinations were invented that expand the characteristics possibility set, our proposition implies that such characteristics combination should be allowed by the second-best optimal tax system. In our context, technology-driven product innovations can be modeled as an outward shift of the GPF. Obviously, such a change of technology would change the underlying parameters of the optimal tax problem and might therefore imply a different optimal tax system, as would be the case in the standard model when technological innovation occurs. Whether or not a tax reform is necessary following such innovations depends on the form of innovation (the change in the GPF). For example, if the innovation is one that proportionally improves all existing goods (a parallel shift of the GPF), relative characteristics and the set of efficient goods remain the same, and so the optimal tax system does not change. By contrast, if an innovation favors specific regions in characteristics space (say, by creating a new local peak on the GPF, or expanding an existing one), it may change the set of socially efficient goods, in which case the tax system would in general have to undergo reform.[27]

Finally, we can state the following about optimal tax rates:

Remark 3 (Optimal Tax Rates) *Whether or not the tax system avoids tax-driven product innovation, optimal tax rates on the two goods in equilibrium satisfy the rule in Proposition 1.*

Proof: From Lemma 1, we always have at most two goods in equilibrium. Conditional on the characteristics combinations of those two goods, the optimal tax problem in eq. (26) is a special case of the problem in eq. (16). □

3.4 The Second-Best Optimal Allocation in the General $N \times M$ Case

The two-characteristic model is illuminating to consider because it allows a simple graphical exposition of the model and results. However, the basic reasoning presented above does not rely on the dimensionality of the problem, and in this section we generalize our results to the case of M characteristics and N market goods chosen from a continuous set of potential goods in M-dimensional characteristics space. Consider therefore a consumer with a utility function $u = (x_0, \mathbf{z})$ where characteristics are generated from market goods according to $\mathbf{z} = \mathbf{Cx}$. Notice that, although we maintain the simplifying assumption from the previous section that leisure x_0 is a direct argument in the utility function, the specification is still flexible enough to retain

[27] An interesting research question would be to explore optimal tax policy when taxes directly affects the incentives for technology-driven product innovations via effects on R&D. As a reduced form, this is a case where the location of the GPF depends directly on the tax system. This question is beyond the scope of this paper and we leave it as a topic for future research.

the standard model $u(\mathbf{x})$ as a special case (when $\mathbf{C}$ is diagonal) and therefore does not impose a specific optimal tax rate structure.

We start by dividing the set of potential goods into those that are located on the no-tax Characteristics Possibility Frontier (socially efficient goods), and those that are located at the interior of this frontier (socially inferior goods). The Characteristics Possibility Frontier represents cost minimization at any given bundle of characteristics $\mathbf{z}$. At no-tax prices, there will be a set of goods $1, ..., N$ solving this problem, and we denote by $\mathbf{C}^*$ the characteristics matrix of these socially efficient goods. Notice that, if a good i is located at the interior of the no-tax Characteristics Possibility Frontier, this means that the characteristics provided by a unit of this good, $\mathbf{c}_i = (c_{1i}, ..., c_{Mi})$, can be achieved at a lower cost under no-tax prices by purchasing non-negative amounts of the N socially efficient goods.

One question is whether a second-best optimal tax system allows for the introduction of a new good that is not part of the socially efficient set $1, ..., N$. We only have to account for one such good, $N+1$, and show that no matter where this good is located within the feasible set, it is dominated by the socially efficient goods in the government's problem. As discussed above, this statement does not imply that no new goods are introduced following the imposition of taxation, because demand-substitution effects will change the bundle of characteristics consumed $\mathbf{z}$, and the set of cost efficient goods is endogenous to the chosen bundle.

Good $N+1$ has characteristics $\mathbf{c}_{N+1} = (c_{1N+1}, ..., c_{MN+1})$, which may be located anywhere on a hyperplane in $\mathbb{R}^M$ of feasible goods, $g(\mathbf{c}_{N+1}) = 0$. We may write the consumption technology as $\mathbf{z} = \mathbf{C}^*\mathbf{x} + \mathbf{c}_{N+1}x_{N+1}$, and we may define the utility function in terms of the distance function as $a(u, x_0, \mathbf{z}) = a(u, x_0, \mathbf{C}^*\mathbf{x} + \mathbf{c}_{N+1}x_{N+1}) = 1$.

The government maximizes utility subject to an exogenous revenue requirement and a zero tax on leisure. We can formulate this problem as maximizing u with respect to $x_0, ..., x_N, x_{N+1}$, subject to

$$(i)\ a(u, x_0, \mathbf{C}^*\mathbf{x} + \mathbf{c}_{N+1}x_{N+1}) = 1,\ (ii)\ 1 - \sum_{i=0}^{N+1} x_i = R,\ (iii)\ a_0(u, x_0, \mathbf{C}^*\mathbf{x} + \mathbf{c}_{N+1}x_{N+1}) = 1, \tag{27}$$

where $\mathbf{c}_{N+1}$ must satisfy $g(\mathbf{c}_{N+1}) = 0$. This problem is analogous to the earlier problem (with a fixed set of goods) in eq. (15). Condition (iii) implicitly defines a function $x_0 =$

$x_0(u, x_1, ..., x_N, x_{N+1})$, and the government therefore maximizes

$$u - \rho\left[a\left(u, x_0\left(.\right), \mathbf{C}^*\mathbf{x} + \mathbf{c}_{N+1}x_{N+1}\right) - 1\right] + \mu\left[1 - x_0\left(.\right) - \sum_{i=1}^{N+1} x_i - R\right] \tag{28}$$

We can then state the following:

Proposition 6 (Product Efficiency) *In the second-best optimal allocation, a socially inferior good such as* $N+1$ *is not introduced. Only goods on the no-tax Characteristics Possibility Frontier* $1, ..., N$ *are allowed, and the set of available goods is therefore first-best.*

Proof: In the appendix. □

This proposition is analogous to Proposition 5, and the interpretations are the same.

4 Conclusion

Any tax system imposing selective commodity taxation must have procedures for assigning different goods to tax rate categories, and real-world tax legislation does this on the basis of observable characteristics. Writing tax laws in terms of characteristics is intuitive and allows a tax system to handle a constantly evolving set of available goods. In this paper, we recast the theory of optimal taxation in the language of characteristics using the Gorman-Lancaster model of consumer behavior, and develop a theory of tax-driven product innovation and optimal line drawing. We present optimal tax rules that depend on characteristics in an intuitive way: the closer two goods are in characteristics space, the greater their substitutability and the smaller the optimal tax rate differential. We point out that, whenever the number of tax instruments is finite, tax systems have to draw lines that define tax-rate regions in characteristics space. Such lines are associated with unavoidable discontinuities, or notches, in tax liability as a function of characteristics, and can create incentives to introduce new goods that are similar to the existing goods as well as qualitatively different goods that—in order to avoid the high tax rate—have characteristics *just* on the low-tax side of the line. All goods that are introduced in response to the tax system are more intensive in the high-tax characteristic than the existing goods in the same tax region. These goods are socially inferior (but privately optimal) in characteristics space, and therefore represents a distortion in the set of available goods.

We show that second-best optimal tax systems draw lines so as to completely avoid such tax-driven product innovations: only goods located on the frontier of the no-tax characteristics possibility set should be allowed in the market. Hence, although the tax system is second-best, the set of goods produced is first-best given the demand for characteristics. This may be seen as a form of production efficiency theorem applied to the production of characteristics from goods; we refer to it as a *product efficiency theorem*. The result does not rule out that the tax system affects the equilibrium set of goods being produced, because the set of socially efficient goods depends on the demand for characteristics, which changes as the result of taxation.

The analysis is framed in terms of the classic Ramsey model of second-best efficient revenue collection, because we feel this is a natural starting point for establishing a characteristics-based theory of optimal taxation and line drawing. In practice, there are of course several other reasons for tax differentiation such as externalities, internalities, or distributional equity. However, whatever the motivation for tax differentiation, a tax system has to assign goods to tax rate categories by drawing lines in characteristics space, which give rise to the kind of effects we have explored in this paper. Hence, we believe that the basic insights presented here are widely applicable, regardless of the reason for tax rate differentiation.

The existing literature does not address the salient features of real-world tax systems that we explore: characteristics-based tax rules, line drawing and notches in characteristics space, and the tax-driven product innovation this generates. Although we have addressed these issues in the context of a consumption tax, they are a ubiquitous feature of all forms of taxation. As an example, consider a Nordic-style dual income tax in which labor income and capital income are subject to separate rate schedules. Any dual income tax must differentiate—i.e., draw a line—between the two types of income, especially with regard to self-employment income. At one time in the Norwegian dual income tax, the complicated rules used to determine whether small business income is taxed as capital or labor income involved notches, such as where shareholders who worked less than 300 hours in the business were not classified as active owners and thus whose income was treated as (preferentially taxed) capital income. Another example arises whenever an income tax allows for the deductibility of interest payments by businesses but not for the return to the providers of equity capital. In this situation, it is inevitable that financial innovation will produce securities that are just debt-like enough to qualify for the deduction. In the U.S., the tax code has not defined debt and equity explicitly, and courts have articulated many factors (i.e., characteristics) to be considered in classification; one well-known case listed sixteen factors to be considered, including the provision of a fixed rate of interest and a provision for redemption at the option of the holder.[28] As a final example, many countries' income tax systems must distinguish between employees, on whose behalf employers must withhold and remit tax liability, and independent contractors, for whom the employer does not have these, and other, legal responsibilities. In the U.S., the IRS uses a 20-factor characteristic-based test that guides the categorization, involving such factors as whether the work is performed on the business's premises and whether the employer has the right to discharge the worker.

Extending the insights of the model to a broader set of such situations will in some cases require some modifications to the model developed here. One challenge will be to provide a principled explanation for why the optimal tax treatment is not uniform, for example between debt and equity finance or among workers with varying relationships with employers. Another will be to allow for heterogeneous tastes with respect to the underlying objects of demand, whether they are the characteristics that goods provide or the stochastic pattern of returns that securities provide, which may generate a "thick" set of goods (or, e.g., securities) in equilibrium. Such enrichments will broaden the policy applicability of the approach we have developed, which can address product innovation, some of which is purely tax-driven, and line drawing.

A Appendix

A.1 Proof of Proposition 1

From eq. (16), the first-order condition with respect to x_j is given by

$$\rho \left[\frac{\partial a}{\partial x_j} + \frac{\partial a}{\partial x_0} \frac{\partial x_0 (.)}{\partial x_j} \right] + \mu \left[\frac{\partial x_0 (.)}{\partial x_j} + q_j \right] = 0, \qquad \forall j.$$

Using the implicit function theorem, we have $\frac{\partial x_0(.)}{\partial x_j} = -\frac{a_{0j}}{a_{00}}$. By inserting this along with $\frac{\partial a}{\partial x_j} = a_j = p_j$, $\frac{\partial a}{\partial x_0} = a_0 = 1$, and $q_j = p_j - t_j$, and subtracting the first-order conditions for goods i and j, we obtain

$$\frac{t_j}{p_j} - \frac{t_i}{p_i} = \frac{\rho + \mu}{\mu a_{00}} \cdot \left\{ \frac{a_{0i}}{a_i} - \frac{a_{0j}}{a_j} \right\}, \qquad \forall j.$$

Using the symmetry of the Antonelli terms and the identity $\frac{a_{i0}}{a_i} - \frac{a_{j0}}{a_j} = \frac{\partial \log(a_i/a_j)}{\partial x_0}$, the above expression is equivalent to the first equality in eq. (17).

[28] Fin Hay Realty v. United States, 398 F.2d 694 (3d Cir. 1968).

To obtain the second equality, we rewrite (11) as

$$\frac{a_i}{a_j} = \sum_{k=0}^{M} \frac{u'_k c_{ki}}{\nabla\mathbf{u}\cdot\mathbf{c}_j} = \sum_{k=0}^{M} \frac{u'_k c_{kj}}{\nabla\mathbf{u}\cdot\mathbf{c}_j}\cdot\frac{c_{ki}}{c_{kj}}.$$

In this expression, all marginal utilities are "compensated" in the sense that consumption levels are divided through by the distance function $a\,(u, \mathbf{Cx})$.[29] By defining $\omega_{kj} \equiv \frac{u'_k c_{kj}}{\nabla\mathbf{u}\cdot\mathbf{c}_j}$, we obtain the second equality in eq. (17). $\square$

A.2 Proof of Lemma 1

The problem of finding the optimal set of goods $1, ..., N$ can be formulated as choosing quantities of goods $x_1, ..., x_N$ at different points on the GPF, $g\,(.)$, so as to determine at which characteristics combinations quantities are positive. This problem can be formulated either as cost minimization at a given characteristics bundle (z_1^*, z_2^*), or as maximizing the amount of one characteristic (such as z_2) at a given amount of the other characteristic ($z_1 = z_1^*$) and at a given total expenditure on goods (y^*), i.e.

$$\max_{x_1,...,x_N} \sum_{i=1}^{N} g\,(c_{1i})\,x_i \quad \text{subject to} \quad \sum_{i=1}^{N} c_{1i}x_i = z_1^*, \quad \sum_{i=1}^{N} x_i = y^*, \quad \text{and} \quad x_i \geq 0 \quad i = 1, ..., N.$$

Denote by (α_1, α_2) the Lagrange multipliers associated with the two equality constraints. Assume that goods 1 and 2 are introduced in the market (i.e., $x_1 > 0$ and $x_2 > 0$) and denote their characteristics vectors by $(c_{11}^*, g\,(c_{11}^*))$ and $(c_{12}^*, g\,(c_{12}^*))$. Then the following first-order conditions must be satisfied

$$g\,(c_{11}^*) + \alpha_1^* c_{11}^* + \alpha_2^* = 0, \tag{29}$$

$$g\,(c_{12}^*) + \alpha_1^* c_{12}^* + \alpha_2^* = 0, \tag{30}$$

where (α_1^*, α_2^*) denote the solution to these two equations.

Under these conditions, will any other good $j \neq 1, 2$ be introduced? This depends on the gain of increasing x_j around $x_j = 0$ given (α_1^*, α_2^*) determined by (29)-(30). Either we have

$$g\,(c_{1j}) + \alpha_1^* c_{1j} + \alpha_2^* \leq 0, \quad \forall j \neq 1, 2, \tag{31}$$

[29] The term "compensated" may not be entirely appropriate here, because this usually refers to a situation where income is changed so as to keep utility constant. Here, we are instead changing the consumption vector by a scale factor in order to keep utility constant.

in which case only goods 1 and 2 are introduced, or we have

$$g\left(c_{1j}\right)+\alpha_1^* c_{1j}+\alpha_2^*>0, \quad \text{for some } j\neq 1,2,$$

in which case j is introduced. In the latter case, one possibility is that good j replaces good 1 or 2 (or both), so that at most two goods are introduced in the market. Alternatively, if good j were to be introduced together with 1 and 2, a necessary condition for an optimum to exist is that

$$g\left(c_{1j}\right)+\alpha_1^* c_{1j}+\alpha_2^*=0,$$

along with conditions (29)-(30). These three conditions imply

$$g\left(c_{1j}\right)+\alpha_1^* c_{1j}=g\left(c_{11}^*\right)+\alpha_1^* c_{11}^*=g\left(c_{12}^*\right)+\alpha_1^* c_{12}^*,$$

which in turn implies that, for any ν, we have

$$\begin{aligned} g\left(c_{1j}\right)+\alpha_1^* c_{1j} &= \nu\left[g\left(c_{11}^*\right)+\alpha_1^* c_{11}^*\right]+(1-\nu)\left[g\left(c_{12}^*\right)+\alpha_1^* c_{12}^*\right] \\ &= \left[\nu g\left(c_{11}^*\right)+(1-\nu)\, g\left(c_{12}^*\right)\right]+\alpha_1^*\left[\nu c_{11}^*+(1-\nu)\, c_{12}^*\right]. \end{aligned} \tag{32}$$

Then consider $\nu^*\equiv\frac{c_{12}^*-c_{1j}}{c_{12}^*-c_{11}^*}$, which implies $c_{1j}=\nu^* c_{11}^*+(1-\nu^*)\, c_{12}^*$. Then eq. (32) at $\nu=\nu^*$ implies that $g\left(c_{1j}\right)=\nu^* g\left(c_{11}^*\right)+(1-\nu^*)\, g\left(c_{12}^*\right)$. Hence, the third good $\left(c_{1j}, g\left(c_{1j}\right)\right)$ is a linear combination in characteristics space of the other two goods, $\left(c_{11}^*, g\left(c_{11}^*\right)\right)$ and $\left(c_{12}^*, g\left(c_{12}^*\right)\right)$. In this case, one of the three goods is spanned by the other two in characteristics space, and can therefore be eliminated without reducing characteristics. Thus, in any case, at most two goods are sufficient to maximize characteristics possibilities.

Moreover, by ordering goods so that goods 1 and 2 are introduced in equilibrium, eqs (29)-(30) yield

$$\alpha_1^* = \frac{g\left(c_{11}^*\right)-g\left(c_{12}^*\right)}{c_{12}^*-c_{11}^*}, \tag{33}$$

$$\alpha_2^* = \frac{g\left(c_{12}^*\right)c_{11}^*-g\left(c_{11}^*\right)c_{12}^*}{c_{12}^*-c_{11}^*}. \tag{34}$$

Because no other good, $j\neq 1,2$, is introduced, the inequality (31) is satisfied. By inserting (33)-(34) into (31), we obtain eq. (22) in the lemma. $\square$

A.3 Proof of Proposition 4

Assume that, in the no-tax equilibrium, both goods 1 and 2 are in the market at $\mathbf{c}_1^* \equiv (c_{11}^*, c_{21}^*)$ and $\mathbf{c}_2^* \equiv (c_{12}^*, c_{22}^*)$. The no-tax CPF is then a straight line from $\mathbf{c}_1^*$ to $\mathbf{c}_2^*$ with a slope (in absolute value) equal to $\frac{c_{22}^* - c_{21}^*}{c_{11}^* - c_{12}^*}$. Because goods 1 and 2 are optimal goods, the no-tax CPF is tangent to the no-tax GPF, $g(.)$, at points $\mathbf{c}_1^*$ and $\mathbf{c}_2^*$. That is, we have

$$g'(c_{11}^*) = g'(c_{12}^*) = \frac{c_{22}^* - c_{21}^*}{c_{11}^* - c_{12}^*}.$$

Consider the tax system in Proposition 4 with $t_2 > t_1$ (if $t_2 < t_1$, the arguments are simply reversed). The after-tax CPF associated with goods 1 and 2 is a straight line through $\frac{\mathbf{c}_1^*}{1+t_1}$ and $\frac{\mathbf{c}_2^*}{1+t_2}$ with a slope (in absolute value) equal to $\frac{\frac{1+t_1}{1+t_2}c_{22}^* - c_{21}^*}{c_{11}^* - \frac{1+t_1}{1+t_2}c_{12}^*}$, which is lower than the original slope given $t_2 > t_1$. The after-tax GPF, $g_t(.)$, relates to the no-tax GPF, $g(.)$, in the following way. If good j at (c_{1j}, c_{2j}) is a point on $g(.)$ and is subject to the tax rate $t_j \in (t_1, t_2)$, then $\left(\frac{c_{1j}}{1+t_j}, \frac{c_{2j}}{1+t_j}\right)$ is a point on $g_t(.)$. This implies that $\left(\frac{c_{1j}}{1+t_j}, \frac{g(c_{1j})}{1+t_j}\right)$ is a point on $g_t(.)$, so that $\frac{g(c_{1j})}{1+t_j} = g_t\left(\frac{c_{1j}}{1+t_j}\right)$. We then have $g'(c_{1j}) = g_t'\left(\frac{c_{1j}}{1+t_j}\right)$.

For part (a) of the proposition, consider good 2 at $\frac{\mathbf{c}_2^*}{1+t_2} = \left(\frac{c_{12}^*}{1+t_2}, \frac{c_{22}^*}{1+t_2}\right)$. As $\mathbf{c}_2^*$ is a point on $g(.)$, $\frac{\mathbf{c}_2^*}{1+t_2}$ is a point on $g_t(.)$. The slope at this point is $g_t'\left(\frac{c_{12}^*}{1+t_2}\right) = g'(c_{12}^*)$. Using that the no-tax CPF is tangent to $g(.)$ at $\mathbf{c}_2^*$, we have

$$g_t'\left(\frac{c_{12}^*}{1+t_2}\right) = g'(c_{12}^*) = \frac{c_{22}^* - c_{21}^*}{c_{11}^* - c_{12}^*} > \frac{\frac{1+t_1}{1+t_2}c_{22}^* - c_{21}^*}{c_{11}^* - \frac{1+t_1}{1+t_2}c_{12}^*}.$$

Hence, the after-tax GPF, $g_t(.)$, is steeper than the after-tax CPF associated with goods 1 and 2 at point $\frac{\mathbf{c}_2^*}{1+t_2}$. Because the GPF and CPF are not tangent at good 2 in the presence of a tax system, this can no longer be an optimal good and will be eliminated. In particular, because the after-tax GPF is steeper than the after-tax CPF at $\frac{\mathbf{c}_2^*}{1+t_2}$, replacing good 2 by a good that provides more of characteristic 2 expands the characteristics possibility set. Hence, if any high-tax good survives in equilibrium, it will contain more of the high-tax characteristic that the original high-tax good.

For part (b), we consider good 1 at $\frac{\mathbf{c}_1^*}{1+t_1} = \left(\frac{c_{11}^*}{1+t_1}, \frac{c_{21}^*}{1+t_1}\right)$. Assume first that $\ell > \frac{g(c_{11}^*)}{c_{11}^*}$, so that good 1 is located in the interior of the low-tax region. Then we can make an argument exactly like the one for good 2 to show that the after-tax GPF, $g_t(.)$, is steeper than the after-tax CPF associated with goods 1 and 2 at point $\frac{\mathbf{c}_1^*}{1+t_1}$. Hence, there exists low-tax goods to the left of good

1 (i.e. with more of the high-tax characteristic) that can expand the characteristics possibility set. If the consumer wants to consume a bundle to the left of good 1, this good will be replaced by one or two new low-tax goods on its left. On the other hand, if $\ell = \frac{g(c_{11}^*)}{c_{11}^*}$, there are no low-tax goods to the left of good 1, and the good will then survive. The proof of the rest of part (b) is straightforward.

A.4 Proof of Proposition 5

To determine whether the government finds it optimal to allow for good 3, we have to consider the gain of increasing x_3 around $x_3 = 0$, given that goods 1 and 2 are in the market at their second-best optimal levels. Good 3 should be introduced if the gain of increasing x_3 (evaluated at $x_3 = 0$) is positive, i.e. if

$$-\rho\left[\frac{\partial a}{\partial x_3} + \frac{\partial a}{\partial x_0}\frac{\partial x_0}{\partial x_3}\right] - \mu\left[\frac{\partial x_0}{\partial x_3} + 1\right] > 0. \tag{35}$$

From the function $x_0 = x_0\left(u, x_1 + c_{13}x_3, x_2 + g\left(c_{13}\right)x_3\right)$, we have

$$\frac{\partial x_0}{\partial x_3} = \frac{\partial x_0}{\partial x_1}c_{13} + \frac{\partial x_0}{\partial x_2}g\left(c_{13}\right),$$

so that (35) can be written as

$$-\rho\frac{\partial a}{\partial x_3} - \mu - \left(\rho\frac{\partial a}{\partial x_0} + \mu\right)\frac{\partial x_0}{\partial x_1}c_{13} - \left(\rho\frac{\partial a}{\partial x_0} + \mu\right)\frac{\partial x_0}{\partial x_2}g\left(c_{13}\right) > 0. \tag{36}$$

This condition must be evaluated at a point where the production efficient goods 1 and 2 are in the market at their second-best optimal levels, which implies

$$-\left(\rho\frac{\partial a}{\partial x_0} + \mu\right)\frac{\partial x_0}{\partial x_1} = \rho\frac{\partial a}{\partial x_1} + \mu, \tag{37}$$

$$-\left(\rho\frac{\partial a}{\partial x_0} + \mu\right)\frac{\partial x_0}{\partial x_2} = \rho\frac{\partial a}{\partial x_2} + \mu. \tag{38}$$

Conditions (36)-(38) may then be combined to give

$$\rho\frac{\partial a}{\partial x_1}c_{13} + \rho\frac{\partial a}{\partial x_2}g\left(c_{13}\right) - \rho\frac{\partial a}{\partial x_3} - \mu + \mu c_{13} + \mu g\left(c_{13}\right) > 0. \tag{39}$$

The functional structure of the distance function $a\left(u, x_0, x_1 + c_{13}x_3, x_2 + g\left(c_{13}\right)x_3\right)$ implies the following relationship between derivatives:

$$\frac{\partial a}{\partial x_3} = \frac{\partial a}{\partial x_1}c_{13} + \frac{\partial a}{\partial x_2}g\left(c_{13}\right),$$

which we insert into (39) to arrive at our final expression

$$g(c_{13}) > 1 - c_{13}.$$

This requirement is precisely ruled out by condition (22) in Lemma 1 under the normalizations $\mathbf{c}_1^* \equiv (1, 0)$ and $\mathbf{c}_2^* \equiv (0, 1)$, which implies that a socially inferior good such as good 3 is characterized by $g(c_{13}) \leq 1 - c_{13}$. $\square$

A.5 Proof of Proposition 6

To determine whether the government finds it optimal to allow for socially inefficient goods, consider the gain of increasing x_{N+1} around $x_{N+1} = 0$. It is optimal to introduce good $N + 1$ only if the gain of increasing x_{N+1} is positive (when evaluated at $x_{N+1} = 0$), i.e.

$$-\rho \left[\frac{\partial a}{\partial x_{N+1}} + \frac{\partial a}{\partial x_0} \frac{\partial x_0}{\partial x_{N+1}} \right] - \mu \left[\frac{\partial x_0}{\partial x_{N+1}} + 1 \right] > 0,$$

which can be written as

$$-\rho \frac{\partial a}{\partial x_{N+1}} - \mu - \frac{\partial x_0}{\partial x_{N+1}} \left[\rho \frac{\partial a}{\partial x_0} + \mu \right] > 0. \tag{40}$$

As good $N + 1$ is a socially inefficient good, the characteristics provided by a unit of this good, $\mathbf{c}_{N+1} = (c_{1N+1}, ..., c_{MN+1})$, can be achieved at a lower cost (under no-tax prices) by a linear combination of socially efficient goods $1, ..., N$. Denote by $\bar{x}_1, ..., \bar{x}_N$, the non-negative amounts of the socially efficient goods required to yield the same characteristics as one unit of $N + 1$. We then have

$$\frac{\partial x_0}{\partial x_{N+1}} = \frac{\partial x_0}{\partial x_1} \bar{x}_1 + ... + \frac{\partial x_0}{\partial x_N} \bar{x}_N,$$

which can be inserted into (40) in order to get

$$-\rho \frac{\partial a}{\partial x_{N+1}} - \mu - \frac{\partial x_0}{\partial x_1} \left[\rho \frac{\partial a}{\partial x_0} + \mu \right] \bar{x}_1 - ... - \frac{\partial x_0}{\partial x_N} \left[\rho \frac{\partial a}{\partial x_0} + \mu \right] \bar{x}_N > 0. \tag{41}$$

This condition must be evaluated at a point where the socially efficient goods $1, ..., N$ are in the market at their second-best optimal levels (which may be zero), so that the potential welfare gains from all those goods have been fully exploited. This implies

$$-\frac{\partial x_0}{\partial x_i} \left[\rho \frac{\partial a}{\partial x_0} + \mu \right] \leq \rho \frac{\partial a}{\partial x_i} + \mu, \quad i = 1, ..., N. \tag{42}$$

Conditions (41) and (42) can be combined to give

$$-\rho\frac{\partial a}{\partial x_{N+1}} - \mu + \left[\rho\frac{\partial a}{\partial x_1} + \mu\right]\bar{x}_1 + ... + \left[\rho\frac{\partial a}{\partial x_N} + \mu\right]\bar{x}_N > 0,$$

where we use that $\bar{x}_1, ..., \bar{x}_N$ are non-negative. We can rewrite this to

$$\rho\frac{\partial a}{\partial x_1}\bar{x}_1 + ... + \rho\frac{\partial a}{\partial x_N}\bar{x}_N - \rho\frac{\partial a}{\partial x_{N+1}} + \mu\bar{x}_1 + ... + \mu\bar{x}_N - \mu > 0. \tag{43}$$

Moreover, because one unit of good $N+1$ corresponds in characteristics to $\bar{x}_1, ..., \bar{x}_N$, we have

$$\frac{\partial a}{\partial x_{N+1}} = \frac{\partial a}{\partial x_1}\bar{x}_1 + ... + \frac{\partial a}{\partial x_N}\bar{x}_N,$$

which we insert into (43) to arrive at our final expression

$$\bar{x}_1 + ... + \bar{x}_N > 1.$$

Recall that, in the no-tax equilibrium, all consumer prices equal 1 and it therefore costs 1 to obtain the characteristics of one unit of good $N + 1$. Alternatively, in the no-tax situation, the above condition implies that it would cost $\bar{x}_1 + ... + \bar{x}_N > 1$ to obtain the exact same characteristics by combining goods $1, ..., N$. This contradicts that goods $1, ..., N$ are the socially efficient goods. $\square$

References

Auerbach, A.J. and J.R. Hines Jr. (2002). "Taxation and Economic Efficiency," in A.J. Auerbach and M. Feldstein (eds.), *Handbook of Public Economics* vol. 3, Elsevier, Amsterdam.

Banks, J. and P. Diamond (2008). "The Base for Direct Taxation." IFS Working Paper, forthcoming in *Reforming the Tax System for the 21st Century: The Mirrlees Review.*

Becker, G.S. (1965). "A Theory of the Allocation of Time." *Economic Journal* 75, 493-517.

Belan, P. and S. Gauthier (2004). "Optimal Commodity Grouping in a Partial Equilibrium Framework." *Economics Letters* 83, 49-54.

Belan, P. and S. Gauthier (2006). "Optimal Indirect Taxation with a Restricted Number of Tax Rates." *Journal of Public Economics* 90, 1201-1213.

Belan, P., S. Gauthier, and G. Laroque (2008). "Optimal Grouping of Commodities for Indirect Taxation." *Journal of Public Economics* 92, 1738-1750.

Chamberlin, E.H. (1953). "The Product as an Economic Variable." *The Quarterly Journal of Economics* 67, 1-29.

Christiansen, V.A. (1984). "Which commodity taxes should supplement the income tax?" *Journal of Public Economics* 24, 195-220.

Crawford, I., M. Keen, and S. Smith (2008). "Value-Added Tax and Excises." IFS Working Paper, forthcoming in *Reforming the Tax System for the 21st Century: The Mirrlees Review.*

Deaton, A. (1979). "The Distance Function in Consumer Behaviour with Application to Index Numbers and Optimal Taxation." *Review of Economic Studies* 46, 391-405.

Deaton, A. (1981). "Optimal Taxes and the Structure of Preferences." *Econometrica* 49, 1245-1260.

Deaton, A. and J. Muellbauer (1980). *Economics and Consumer Behaviour*, Cambridge University Press, Cambridge.

Diamond, P.A. and J.A. Mirrlees (1971). "Optimal Taxation and Public Production, Part I: Production Efficiency, Part II: Tax Rules." *American Economic Review* 61, 8-27, 261-278.

Gordon, J.P.F. (1989). "Tax Reform via Commodity Grouping." *Journal of Public Economics* 39, 67–81.

Gorman, W.M. (1970). "Quasi-separable preferences, costs, and technologies." Mimeo, University of North Carolina, Chapel Hill.

Gorman, W.M. (1976). "Tricks with utility functions," in M. Artis and R. Nobay (eds.), *Essays in Economic Analysis*, Cambridge University Press, Cambridge.

Gorman, W.M. (1980). "A Possible Procedure for Analysing Quality Differentials in the Egg Market." *Review of Economic Studies* 47, 843–856.

Harberger, A.C. (1995). "Tax Lore for Budding Reformers," in R. Dornbusch and S. Edwards (eds.), *Reform, Recovery and Growth*, University of Chicago Press, Chicago.

Kaplow, L. (2008). *The Theory of Taxation and Public Economics*, Princeton University Press, Princeton.

Kleven, H.J. (2000). "Optimum Taxation and the Allocation of Time." EPRU Working Paper 2000-16, October 2000.

Kleven, H.J. (2004). "Optimum Taxation and the Allocation of Time." *Journal of Public Economics* 88, 545-557.

Lancaster, K.J. (1966). "A New Approach to Consumer Theory." *Journal of Political Economy* 74, 132-157.

Lancaster, K.J. (1975). "Socially Optimal Product Differentiation." *American Economic Re-*

view 65, 567-585.

Saez, E. (2002). "The Desirability of Commodity Taxation Under Non-Linear Income Taxation and Heterogeneous Tastes." *Journal of Public Economics* 83, 217-230.

Saez, E. (2004). "The Optimal Treatment of Tax Expenditures" *Journal of Public Economics* 88, 2657-2684.

Salanié, B. (2003). *The Economics of Taxation*, MIT Press, Cambridge.

Shaw, J., J. Slemrod, and J. Whiting (2008). "Administration and Compliance." IFS Working Paper, forthcoming in *Reforming the Tax System for the 21st Century: The Mirrlees Review*.

Slemrod, J. and S. Yitzhaki (2002). "Tax Avoidance, Evasion and Administration," in A.J. Auerbach and M. Feldstein (eds.), *Handbook of Public Economics* Vol. 3, Elsevier, Amsterdam.

Sørensen, P.B. (2007). "The theory of optimal taxation: what is the policy relevance?" *International Tax and Public Finance* 14, 383-406.

Weisbach, D.A. (1999). "Line Drawing, Doctrine, and Efficiency in the Tax Law." *Cornell Law Review* 84, 1627-1681.

Weisbach, D.A. (2000). "An Efficiency Analysis of Line Drawing in the Tax Law." *Journal of Legal Studies* 29, 71-97.

Wilson, J.D. (1989). "On the Optimal Base for Commodity Taxation." *American Economic Review* 79, 1196-1206.

Yitzhaki, S. (1979). "A Note on Optimal Taxation and Administrative Cost." *American Economic Review* 69, 475-480.

Figure 1: The Goods Possibility Frontier

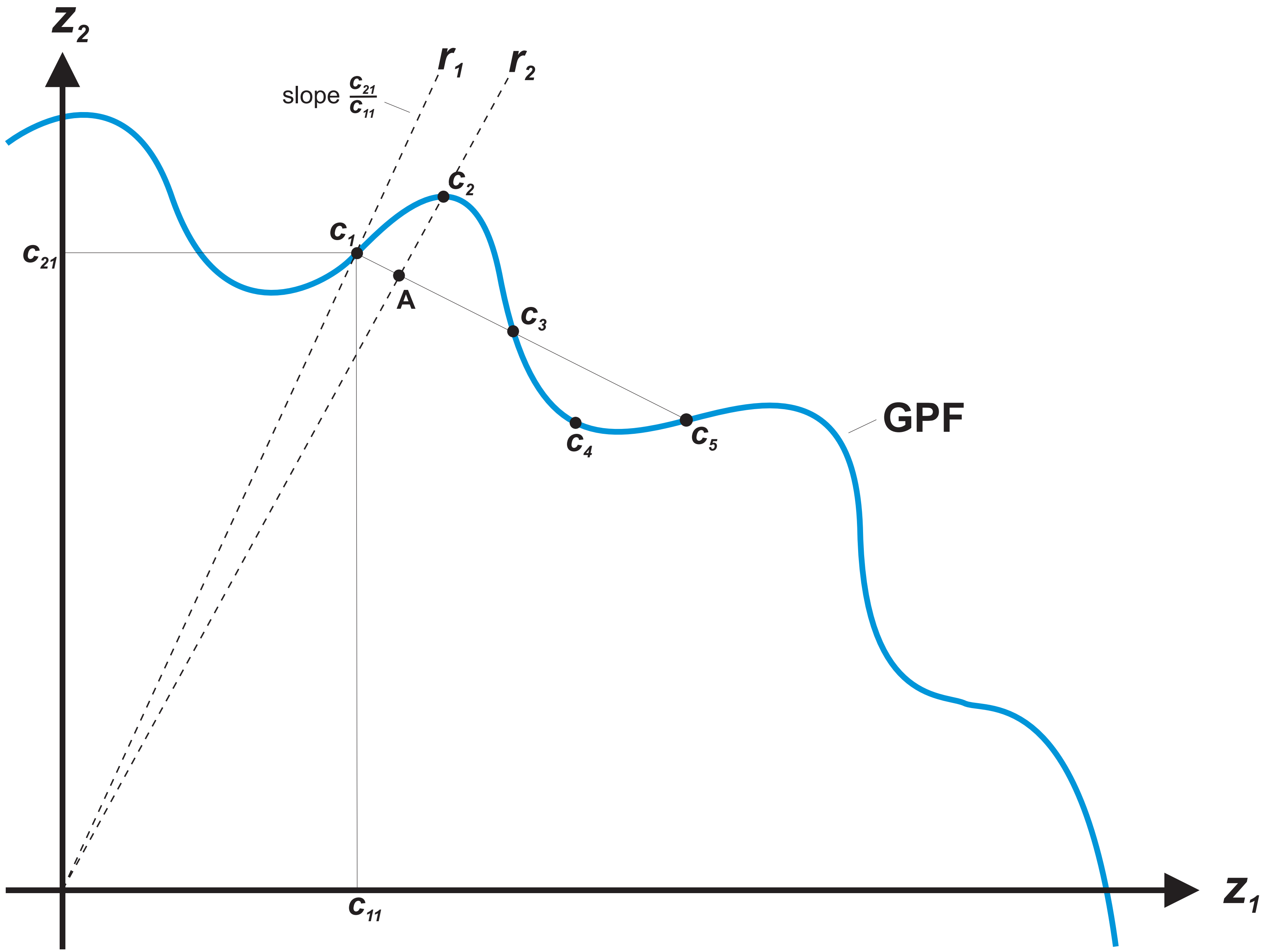

Figure 2: The Characteristics Possibility Frontier

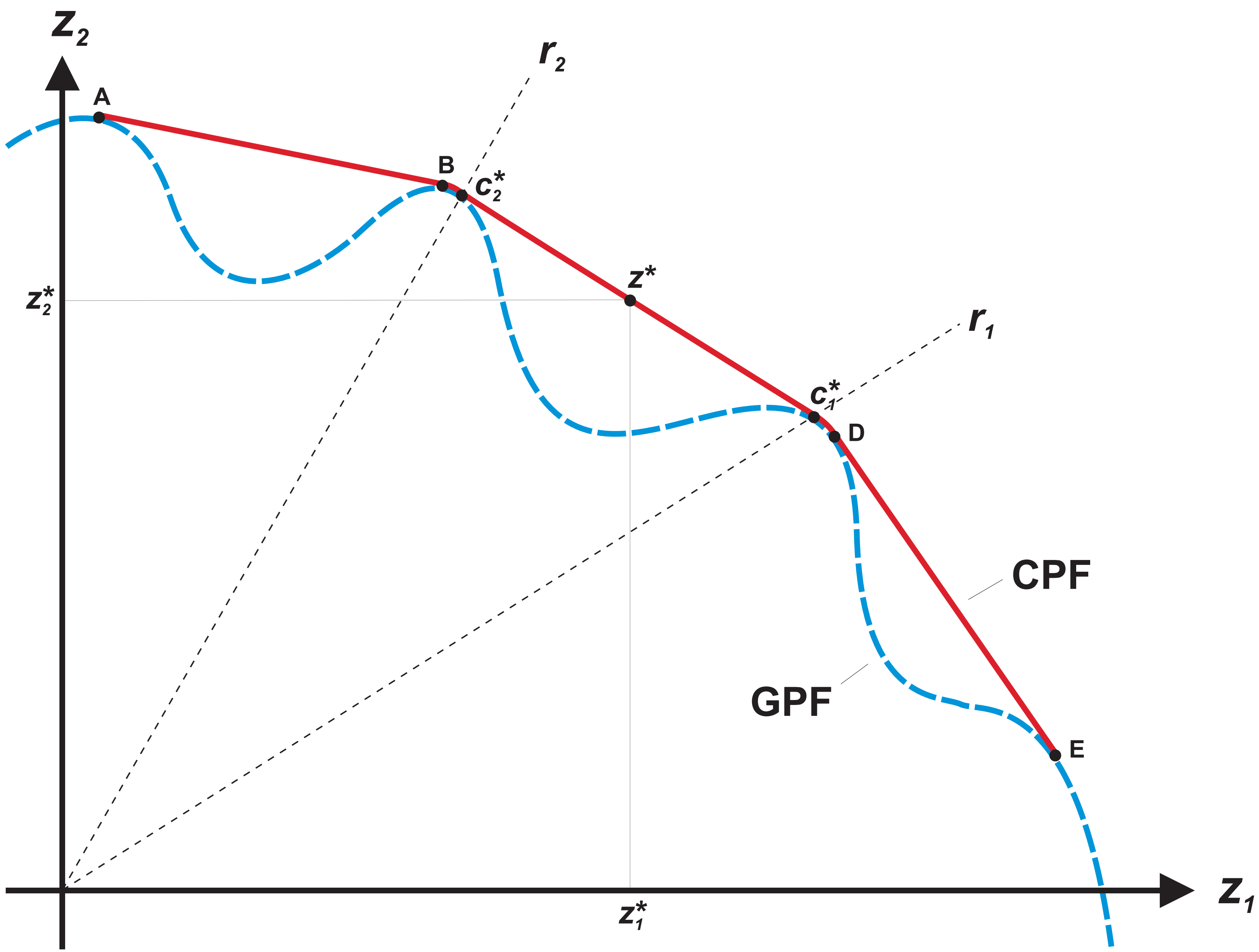

Figure 3: The Impact of Taxation and Line-Drawing on the Goods Possibility Frontier

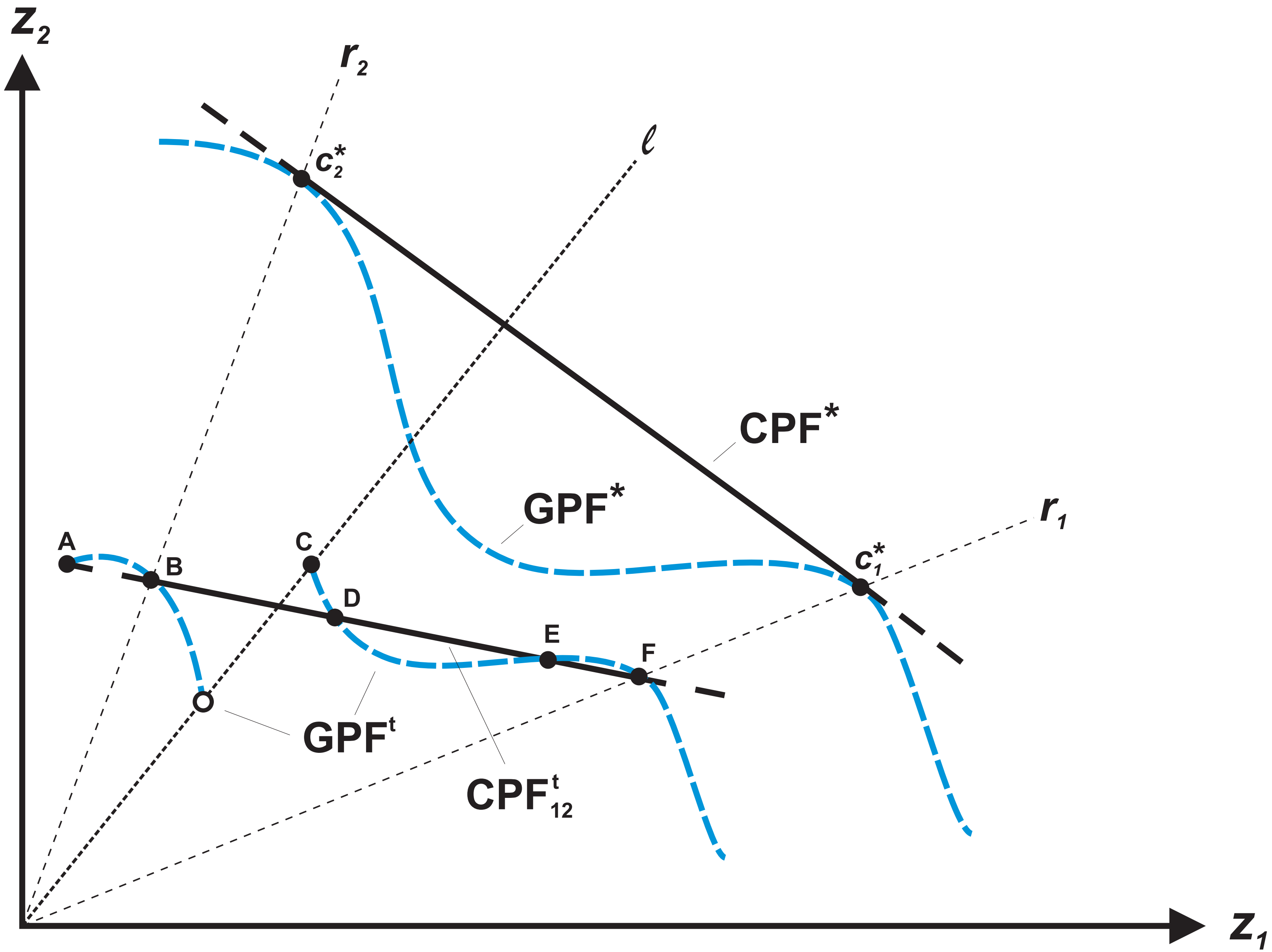

Figure 4: The Impact of Taxation and Line-Drawing on the Characteristics Possibility Frontier